

STIC Search Report

Biotech-Chem Library



STIC Database Tracking Number 1935541

TO: Nancy Vogel
Art Unit: 1636
Location: REM-2A70&2C70
Serial Number: 10/689200
Wednesday, May 25, 2005

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

Protein Sequence Searches – February 2005

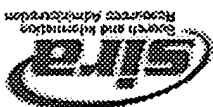
All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension .rnp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.



This Page Blank (uspto)

153541

From:
Sent:
To:

Vogel, Nancy
Monday, May 16, 2005 2:46 PM
STIC-Biotech/ChemLib

Subject: sequence search 10/689,200

please do a sequence search of SEQ ID NO: 1 and 2

of 10/689,200 and return results to me on paper asap.

Thanks,

Nancy Vogel
Patent Examiner
Art Unit 1636
Office: Remson 2A70
Mail Box: Remson 2C70
(571) 272-0780

STAFF USE ONLY

Searcher:
Searcher Phone: 2-
Date Searcher Picked up:
Date Completed:
Searcher Prep/Rev. Time:
Online Time:

Type of Search
NA#: _____
Inference: _____
SPDI: _____
S/L: _____
Oligomer: _____
Encode/Transl: _____
Structure#: _____
Text: _____
Inventor: _____
Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

Date completed: _____
Searcher: Beverly e 2528
Terminal time: _____
Elapsed time: _____
PU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site
STIC
CM-1
Pre-S
Type of Search
N.A. Sequence
A.A. Sequence
Structure
Bibliographic

Vendors
IG
STN
Dialog
APS
Geninfo
SDC
DARC/Questel
Other CCN

This Page Blank (uspto)

Vogel, N.
107689200 Page 1
Seq. IDs 172

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 20, 2005, 13:18:48 ; Search time 6175 Seconds
(without alignments)
10923.022 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaaacacatcatagatc.....ccgycgcaagtcgtgat 1392

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

GenBml:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	628.8	45.2	11493	1 AF283498	AF283498 Azospirillum
C 2	555.6	39.9	300883	1 AE016809	AE016809 Vibrio vulnificans
C 3	541.2	38.9	248850	1 AP005348	AP005348 Vibrio vulnificans
C 4	516.2	37.1	302100	1 AP005087	AP005087 Vibrio vulnificans
C 5	431	31.0	311050	1 BX294133	BX294133 Pirella
C 6	171.2	12.3	110000	1 AE017282	AE017282 of
C 7	133	9.6	10676	1 AE008293	AE008293 Agrobacterium
C 8	133	9.6	10707	1 AE009321	AE009321 Agrobacterium
C 9	122.8	8.8	348077	1 AP003000	AP003000 Mesorhizobium
C 10	119.8	8.6	300150	1 AP005945	AP005945 Bradyrhizobium
C 11	118.2	8.5	110000	1 RME591985	RME591985 of
C 12	116	8.3	196050	1 AL646058	AL646058 Ralstonia
C 13	113.2	8.1	298700	1 AP005951	AP005951 Bradyrhizobium
C 14	109.6	7.9	12129	1 AE007223	AE007223 Sinorhizobium
C 15	104.8	7.5	299750	1 AP006575	AP006575 Gloeobacterium
C 16	104.4	7.5	339650	1 AP003583	AP003583 Nostoc sp
C 17	100.6	7.2	348068	1 BX572604	BX572604 Rhodospirillum rubrum
C 18	98.4	7.1	300700	1 AP006573	AP006573 Gloeobacterium
C 19	98.2	7.1	110000	1 BX950851	BX950851 of

20	95.2	6.8	17155	1 AY273169	AY273169 Rhodospirillum rubrum
21	95	6.8	4148	1 KPNASPEC	KPNASPEC Klebsiella pneumoniae
22	93.4	6.7	13551	1 D90906	D90906 Synechocystis
23	89.8	6.5	3253	1 SYONIRNRT	SYONIRNRT Synechococcus
24	89.2	6.4	300363	1 AE016781	AE016781 Pseudomonas
25	87	6.2	10798	1 AE004604	AE004604 Pseudomonas
26	86	6.2	300300	1 AP005373	AP005373 Thermosyn
27	84.6	6.1	349142	1 BX572599	BX572599 Rhodospirillum rubrum
28	82.4	5.9	1371	6 AR386463	AR386463 Sequence
29	82.4	5.9	299650	1 AP005955	AP005955 Bradyrhizobium
30	80.2	5.8	4080	1 SYOCMP	SYOCMP Synechococcus
31	79.4	5.7	298750	1 AP005375	AP005375 Thermosyn
32	77.8	5.6	302550	1 AP006581	AP006581 Gloeobacterium
33	75.2	5.4	11318	1 AE008005	AE008005 Agrobacterium
34	75.2	5.4	11330	1 AE009039	AE009039 Agrobacterium
35	74.6	5.4	6449	1 PLNRTABC	PLNRTABC Pseudomonas
36	74.4	5.3	347400	1 AP003591	AP003591 Nostoc sp
37	73.8	5.3	7520	1 AE005736	AE005736 Caulobacter
38	72	5.2	132106	1 SYCSLH	SYCSLH Synechocystis
39	71.8	5.2	310325	1 AE016864	AE016864 Pseudomonas
40	70	5.0	3218	1 AE001333	AE001333 Synechococcus
41	69.2	5.0	11321	1 AE005737	AE005737 Caulobacter
42	65.8	4.7	3719	1 AY283674	AY283674 Cyanobacterium
43	63.4	4.6	2210	1 AYNASST	AYNASST X81602 A. vinelandii
44	63.4	4.6	4080	1 SPCNRT	SPCNRT X61625 Synechococcus
45	60.8	4.4	110000	1 CP000011_11	Continuation (12 o

ALIGNMENTS

RESULT 1	AF283498	11493 bp	DNA	linear	BCT 02-JUN-2002
LOCUS	AF283498				
DEFINITION	Azospirillum brasilense nitrate assimilation gene cluster, complete sequence.				
ACCESSION	AF283498				
VERSION	AF283498.1	GI:21666012			
KEYWORDS					
SOURCE					
ORGANISM	Azospirillum brasilense				
REFERENCE					
AUTHORS	Steenhoudt, O., Placet, D., Verreth, C. and Vanderleyden, J.				
TITLE	Nitrate assimilation in Azospirillum brasilense Sp245: mutational and expression analysis of the nas gene cluster				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 11493)				
AUTHORS	Steenhoudt, O.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-JUN-2000) F.A. Janssens Laboratory of Genetics, Catholic University of Leuven, Kard. Mercierlaan 92, Heverlee 3001, Belgium				
FEATURES					
source	location/Qualifiers				
	1..11493				
	/organism="Azospirillum brasilense"				
	/mol_type="genomic DNA"				
	/strain="Sp245"				
	/db_xref="taxon:192"				
gene	complement (526..2190)				
	/gene="nasD"				
CDS	complement (526..2190)				
	/gene="nasD"				
	/codon_start=1				
	/transl_table=11				
	/product="nitrate transporter ATPase"				
	/protein_id="FAM73541.1"				
	/db_xref="GI:21666013"				
	/translation="MALIKNAYKSGVGVNVRGIDLSIAGEFVAIVGSGAKTTLVNMGGLAPDAGIVLFGAPVKGGAERGVFQNYSLMPHLSVDNVALAYAHK GKSPRGVITRAAVTVGATDTRPELSGGMQRQGFARFALMSPEMLLDEL SALDALTRAKLODEIRAIWRKDKTKTVLLITNDVDAIILLADRIIPLTPGCAVTLGPAPF				

```

AVDLPRDRRTAVNHDPDFKRLAEVTAMLMESAGSRAAAVADESITMPDAPIITAP
PKRAYLAGOGGSLDDRYVERFSRYTPATPKGPIIVDGFOLKXMKKGFPIISGSC
CKKSTVLSMVAGLADYSSGGIYLDGEVVGAGDRAVYQASLFPWILTRLENMIGV
DRYVPOAGGERNDIARYLISRGVGNAMHAKTAEISNKKQGVJARAFAFKILL
LDPEFGMLDLSLRMELQEVLMEMWARTQVTAICTSDVDVBAIILADRYVMSGPNAR
IGHILEVDIPIHPTROALLEHPPRYAYREELINFLGCHGTGKAA"
gene
CDS
    /gene="nase"
    complement (2204 . 3295)
    /gene="nase"
    /codon_start=1
    /transl_table=11
    /product="nitrate transporter integral membrane protein"
    /protein_id="AAW73542.1"
    /db_xref="GI:2166014"
    /translation="MTSVTEASTLSSAEMARRKRLAHALNRTAIIQIVGLMIS
    PLIRMAAGSPPROGQELMOQGVPLTALMLFLAAMWMLAPQVNTSIGIIPGAQVWE
    QIQLHADHKAKEAFYDRQAKNAELADPAEYKTRPAKPYLQIITSL
    KYTFEGLGALVAAPLGVACGLSRIVNANMPLIQIFKPSPLAMPLIVTVSASIV
    SSDNPREKSESLTSANTVTLCSLMTPLINTAVGVSSIDDKLMVGVGLDSGTWRR
    LIDPSALPTIFGLRSLDGVGMVTLIAEMLQNPGLGKFWDEFQNGSSSLIARIV
    AVFTIGLIGFLDRVMAFOAAVSHGTG"
gene
CDS
    /gene="nase"
    complement (3455 . 4843)
    /gene="nase"
    /codon_start=1
    /transl_table=11
    /product="nitrate transporter substrate-binding protein"
    /protein_id="AAW73543.1"
    /db_xref="GI:2166015"
    /translation="MIRLPASPTPLLSAATLTALMGSNOAAPLIDVEDOKLTGRK
    LIDMALIAAEKGFEDBSLSTYTLPEQNAKYLLDRVSGELDGHMAGPGLGRTI
    GGTQANVYAFSMDLNGGILTSNEVWRMKNLPKGPDPKPLHDKADLKPAVQ
    YRABGKPFMTGWFVSTHNYELRYLAAGINPGYANDVSGQIOADALISVPP
    QMPATLEAGTIFGVSVGEPMNOQAVKKGIVPIDTEIMKNNPEKVFETGMAKN
    PKTHALVYKALIRAAMLDENGANRAEAKVILAKSEYVGAADKVIANTGTPEYEX
    GKRAVPDPNVPFRNATYTPYSDATWYLTQKRMQGLAEKPDANYDSTAKRYEYK
    TYKARLLVEBKAKADPMTSDGKFLDNGFIDGIAIDGKKEVEYITKLPITGKG
    GQAVQGQQLVG"
gene
CDS
    /gene="nase"
    5514 . 7979
    5514 . 7979
    /gene="nase"
    /codon_start=1
    /transl_table=11
    /product="assimilatory nitrite reductase large subunit"
    /protein_id="AAW73544.1"
    /db_xref="GI:2166016"
    /translation="NSTHTNPPTRPRRLVYVNGNAGITRLEBLAKAPDRIDITYG
    ABPHPNYRIMLSPVLAGKTEPQIVLNGRDWEANGIILTGEDREVIDRANCVTA
    IEGLTVPYKRLIATGSTPVIIPVPGSLTGVGFADLADVDMLEAARGRAVVG
    GGLTGLEANGKLVKGMADVTLIMDLTMRQLDRSAGALRHELEKRGITVLTGADT
    AEIVSEKYSAYLRKNGOELPADLVMAVGIRPNTLGAAGACGSGI:QVDDAMTS
    DAILISVEGCVHKGOTGVVAPLPEMAVVAARLAGSAEAYTGSVSTKIKYGV
    VVSADPFTGKDCEDIVFRDARGVYKRVVKNRILGAVLTGDTROGMYTOMLKG
    TDEAVARDTLIFGOSFGGGAANPKAAVALPDSAEI GCGNGVCKGTIVATITKEL
    SLDVEVAHAKASGSCCTGQVEQLALTLTGDYAEBAKAPKCTCTDITBEVRAI
    TALDKTIPDVQRLMEWTPDGHCRPALNTYLLCEMFGYKDSRSRYIKERVYAN
    IOKDGTYSVPRMNGGLTSKELRAIADVDKKALPTVAVTGGORIDLGVKKEDIPA
    VWADLNAGMVSGHAYAKGLRTYKTCVSEKCEMCFRGNODTGTGVLCEMTKTPH
    VKLAVSGCPKRCABATIKDQGVYCVDSGEIHLVGNAGNHAVALCDLIVETBAEYK
    YTGAVQLRREARVLERTAPWVERVGLDHIRRLVDAEGRALNARLIFSOSFSD
    DWAERANGSVRHEPAPLAQV"
gene
CDS
    /gene="nase"
    7986 . 8342
    7986 . 8342
    /gene="nase"
    /codon_start=1
    /transl_table=11
    /product="assimilatory nitrite reductase small subunit"
    /protein_id="AAW73545.1"
    /db_xref="GI:2166017"

```

```

/translation="MDOSISWITVGSVEDIPQGRSVVRGFDIALFRTVADEVYAL
EPRCHAGGSLSSQIVGHGARVTCPLHSWIDLASBAVGPDBGCAGHCVRVVDVQ
LGLAAPRGCGCGHG"
gene
CDS
    /gene="nase"
    8473 . 11130
    8473 . 11130
    /gene="nase"
    /codon_start=1
    /transl_table=11
    /product="assimilatory nitrate reductase catalytic
    subunit"
    /protein_id="AAW73546.1"
    /db_xref="GI:2166018"
    /translation="MTVPTTQPCVGGVGLVASYSDQSTGVAGDPHPANRKLCS
    KETNLPDTTIGLERLHPVDGRRAVMDAIEVAKRLSDITAKGPPDVAAYVSQI
    LVEDYVAVVKLAKGFLGTANIDTNSRLCAASSVAHGKRGFDADVAIVYDIDLEADV
    LIVGNSLMLKFLINQRLANLKGITGTRLVTDPRTPAADADRHLP:RPGDIDLEALF
    NGLVYLDNHLADNGLIDARVAHHTGALALBAARADASPAAVACGLPLAEV
    EAPYAEVATRRVTIVYSGVNSOGSDTVNAILNVHFLNGRIGAPGMPSPVSGP
    NAMGRBVUGLANOAAHMGFPPEVDRLRRFMNAPRAVDKRGKXAVD:FAVEGCTV
    RAVWATNPVAVSMFPAKVRPALKCEVTVISDCAVDITRLSHIRLPAAGMSKD
    GVTVNSDRTISQORAEPPLAGARPDMWITVQVARAMGPAALFSETPADVREHAIL
    SAFENGSTRAFDIGLAGLSDAEFTLAPWPRRAGSPTRRLPADGFFPEDGRAR
    FVPAVSRLP:RSEVSGSLMLNTGRLDQWHTRTGLSPRLSAHAPECLIDHPDPA
    ARGIDGALVLSSTPVGTALRLRYTEAQVGEAFLPMWTRPFTASCVGRVLDLA
    VDPVSGQDPLKMPVTLTPYPAAMTGFILAEPPVPAHGVMSRPAVPGCH:IELASE
    DLAPBETIAFGPAGATVDYSDAAGVLRANVTHDRLEACFLSTSGALPBGMLIVE
    LLSARLDGAAARRALLSGSPVPTADEGKTVSCGAVGNRLIAI:QSGGLTVVDVG
    MLRAGTNGSGCIPELKEILDAPORSHYA"
ORIGIN
Query Match 45.2%; Score 628.8; DB 1; Length 11493;
Best Local Similarity 68.7%; Pred. No. 8.5e-111;
Matches 912; Conservative 0; Mismatches 407; Indels 9; Gaps 3;
QY 69 GGGTTGACGATTGCGCCCGATGTCGGGGCAGTTGGCAAGCTGAAAGAAATTGAA 128
DB 4787 GGGCTTGAATGCTCGGGTCCGCCAGCGCGCTCGCTGACAGTGGAGAAGACAGCTCAA 4728
QY 129 ATTGGGCTTCATCAAGCTCAGCGACATGAGCGCGCTGGCGGCGCGGAAAAAGGCTT 188
DB 4727 GCTGGGCTTCATCAAGCTCAGCGATATGAGCGCGCGCTCGCATCGCGCGGAGAGGCGCTT 4668
QY 189 CTTTGAAGACAGAGGCGCTTGTGTGCAACTGGAAGCGGCGCACTGGAAGGTGTGAT 248
DB 4667 CTTTGAAGACAGAGGCGCTTGTGTGCAACTGGAAGCGGCGCACTGGAAGGTGTGCT 4608
QY 249 GGAATGAGGCGTGAATGCGAACTGGAAGGCTCGACATGCTGAGCGCGCGCGCTTACG 308
DB 4607 CAGCGGTGATCTCGGGGAGCTGAGCGGCGCGCCACATGCTGGCGCGGACCGCTGGAG 4548
QY 309 GGCAGCGTTGAGCTTCGCGACCAAGGCGCATATCGAGTGCCTTCAGATGGGCTTCAA 368
DB 4547 GGCAGCATCGGCTTCGCGACCAAGGCGCATATCGAGTGCCTTCAGATGGGCTTCAA 4488
QY 369 CGGCAACGCGATTACGCTTCCATGAATTCGGCATCATGATGAAGCCGAATACCC-- 425
DB 4487 CGGCAACGCGATCACCTCTCCACAGAGGTGTGGAGCGCGAAGCCGAACCTGCGCAA 4428
QY 426 GCTGGAAGCGGTAACCGGTCATCCGATTCAGAGGAGATTATCTAAACCGGTGTGCA 485
DB 4427 GGGGCGGAGCGGAGAGCGGCTGACCCGATTCAGAGGCGGCGCTGAAACCGGCTATCGC 4368
QY 486 AAAATTAAGGCGGAAGGCAAGCGGCTTCATATGCGATGACCTTCCGCGCGGATTCGA 545
DB 4367 CCAAGTACCGGGAGAGAGGCAAGCGGCTTCATTCAGATGGGCAATGCTTCCCGGTCCACCA 4308
QY 546 CAACATCAAGCTGCTTACGCTGCGGCGGCTGCGGATCAATCTCGGCTACTATTGCGC 605
DB 4307 CAATTACGAGCTGCGCTATGCTTGGCGCGCGGCGGATCAACCCGCGGCTACTA---CGC 4251
QY 606 GCGGAGGACATTTCCGCGCAATCGGCGCAACGCTTGTGTGCGGTACACCCCGCGGCGC 665

```

Db 4250 GCCCAAGACGCTGCGGCGCAAGATCCAGGCGGACGCGCTGCTCCGTCAACCCCGCCG 4191
 Qy 666 GCAATGCGCGTCCACGCTGAGACCGGACCAATTTTGGTTATTCGCTCGGCGAGCCGTG 725
 Db 4190 GCAGATGCGCGCGACGCTGAGGCGCGGACCAATTTTGGTTATTCGCTCGGCGAGCCGTG 4131
 Qy 726 GAAACGACGCGCTGTTCAAGGCGATAGCGCTGCGGTATTCACCGATGAAGACTCTG 785
 Db 4130 GAACCGACGCGCTGATGAAGGCGATGCGCGGTATTCACCGATGAAGACTCTG 4071
 Qy 786 GAAGGACGCGCGGAAAAAGTCTTTCGCGGTGACCAAAACAAATGCGCGGAGAAATACCCCA 845
 Db 4070 GAAGGACGCGCGGAAAAAGTCTTTCGCGGTGACCAAAACAAATGCGCGGAGAAATACCCCA 4011
 Qy 846 CACTTATCTGCGCGTGAACCAAGCGCTGATTCGCGCGCGCATCTGCTGAGCGCGGACAA 905
 Db 4010 GACCCATCTGCGCGCTGGTCAAGCGCTGATTCGCGCGCGCATCTGCTGAGCGCGGACAA 3951
 Qy 906 TAAAGAAACCGCAAGAGCCATCGAAATGCTGCGCGCAAAAACATACGTCGCTGTA 965
 Db 3950 CAACGCAACCGCGCGGAGCGGTGAAGATCTGCGCAAAATGCTGCGGCGCGCA 3891
 Qy 966 CGTGAAGTCTGCGCGGTGACCAAGCGCGCTTTCGATTAAGAAAGAGCATTAACG 1025
 Db 3890 CGCAAGGTCATGCGCAATCGATGACCGGACCTTTCGATTAAGAAAGAGCATTAACG 3831
 Qy 1026 CGCGCTACCGGACCTTCAACACTTCTTTCGCGCGCGCGCATCTGCTGACAGCAG 1085
 Db 3830 GGGGATTCGCGACTTCAAGCTTCTTTCGCGCTACCAAGCGCATCTGCTGACAGCAG 3771
 Qy 1086 TCGAGTCTGATATCTGACCAAGCTGAGCGCTGCGCGCATGATCAATGAATTCAAACCGGA 1145
 Db 3770 CGCGCTGATCTGATCTGACCAAGCTGAGCGCTGCGCGCATGATCAATGAATTCAAACCGGA 3711
 Qy 1146 CAACGTGATCTGATCTGACCAAGCTGATCTGCGCGCGCATCTGCTGACAGCAG 1205
 Db 3710 CGCGCTGATGACCAAGCTGATCTGCGCGCGCATCTGCTGACAGCAG 3651
 Qy 1206 CAAGGACTGCTGCGCAAGAGCGCAAGCGCGCAAGCTTCCC--TSCCGATACCTC 1262
 Db 3650 GCGGCTGCTGCTGAGAGAGCGCAAGCGCGCAAGCTTCCCCTCGAGCGAGCGG 3591
 Qy 1263 GATCAAGCGCTGCGCAAGCTTCTTCAATCGACAAAGTCCGCTGATGCGCAAGCGCA 1322
 Db 3590 CTACAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3531
 Qy 1323 CCAATTCTCTGCGCAAGCTTCTTCAATCGACAAAGTCCGCTGATGCGCAAGCGCA 1382
 Db 3530 CGAGTACTGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3471
 Qy 1383 GGTCTGG 1390
 Db 3470 ACTGCTG 3463

RESULT 2
 AE016809/c
 LOCUS 300883 bp DNA linear BCT 29-SEP-2003
 DEFINITION Vibrio vulnificus CMCP6 chromosome II section 2 of 6 of the
 complete sequence.
 ACCESSION AE016809 AE016796
 .VERSION AE016809.1 GI:27358288
 .KEYWORDS
 SOURCE
 ORGANISM
 Vibrio vulnificus CMCP6
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.
 1 (bases 1 to 300883)
 Kim, Y.R., Lee, S.E., Kim, C.M., Kim, S.Y., Shin, E.K., Shin, D.H.,
 Chung, S.S., Choy, H.B., Progulskis-Fox, A., Hillman, J.D., Handfield, M.
 and Rhee, J.H.
 Characterization and Pathogenic Significance of Vibrio vulnificus
 Antigens Preferentially Expressed in Septicemic Patients

JOURNAL Infect. Immun. 71 (10), 5461-5471 (2003)
 PUBMED 14500463
 REFERENCE 2 (bases 1 to 300883)
 AUTHORS Jeong, H., Moon, Y.H., and Kim, J.J.
 TITLE Direct Submission
 JOURNAL Submitted (13-DEC-2002) Genotech Corp., 461-6, Jeomun-dong,
 Yuseong-gu, Daejeon 305-811, South Korea
 REFERENCE 3 (bases 1 to 300883)
 AUTHORS Rhee, J.H., Kim, S.Y., Chung, S.S., Lee, S.E., and Choy, H.B.
 TITLE Direct Submission
 JOURNAL Submitted (13-DEC-2002) Department of Microbiology, Genome Research
 Center for Enteropathogenic Bacteria, Chonnam National University
 Medical School, Hak-1-Dong, Dong-Gu, Kwang-Ju 501-746, South Korea

FEATURES
 source
 1. 300883
 /organism="Vibrio vulnificus CMCP6"
 /mol_type="genomic DNA"
 /strain="CMCP6"
 /db_xref="taxon:216895"
 /chromosome="II"
 269. 577
 /locus_tag="VV20274"
 269. 577
 /locus_tag="VV20274"
 /codon_start=1
 /transl_table=1
 /product="Unknown"
 /protein_id="AA007239.1"
 /db_xref="GI:27358289"
 /translation="MILENKPNTSRPFLISNNIKQDFPIYFHEKTPQIIKAKTNSIV
 TISGRKDFIKIGGVBLTTQGRALIGICORALOEKDFVDASLMIADVEKEARCV
 "

gene
 565. 2553
 /locus_tag="VV20275"
 565. 2553
 /locus_tag="VV20275"
 /note="COG2199"
 /codon_start=1
 /transl_table=1
 /product="P08: GDEF domain"
 /protein_id="AA007240.1"
 /db_xref="GI:27358290"
 /translation="MCSLSLSIKNLLPILFGLIFLIAQFHFMSYEKQDQIQ
 NHRISLIAGAAQNLQASVLFNDQKMOEILSFSDPILIVELSLLEGVARVOR
 BMPVNDNVLSPYPERETFYSORIDVEONAIALSTASTQTLQOMQVRLRESV
 VPLAILLSGSTLYLVOKVVIDPVYKLMSMODIVGRILKGLSVSHSDEGELVN
 ARNMQRLALRDSQVHTLDRLEHKAATDVTEAVQTRLLVYNEHGDILHNNAR
 MLFQSDAQRIDQKLSLMCKEFSVANALTSKNLEKLVLCRLPTGSKLILKIST
 RLISKRATLFSIDVTEIETALNRORLAGVEENQDGLVVDHNGLLTMTPAVTR
 MGYPOEILIGKSLKAFPMQIOTIMPTIOESLBYGQGEVLKHYNGNLIPMA
 KVSRLSKVGVQVEFPMVILISDLSGVKEMERLEYLAHHDALGLANRAKGLRHLDE
 LKICNFTHSERFALLVMDLDGFKPINTPRGHDADHYLKYVRELLNEVSGDILVARS
 GDEFVLIRADNHSVTLISRLISLRQPIETLQNTQVQSGIGVAVGPQEKMER
 ILKSADTAMVYKAKSGGQAILIGSEVGACLSDVTS"

gene
 complement (2838..4703)
 /locus_tag="VV20276"
 complement (2838..4703)
 /locus_tag="VV20276"
 /note="similar to mostly Fe transport; COG1629"
 /note="similar to mostly Fe transport; COG1629"
 /codon_start=1
 /transl_table=1
 /product="Outer membrane receptor protein"
 /protein_id="AA007241.1"
 /db_xref="GI:27358291"
 /translation="MEGRVKLLVDAQSGPQVFDGPIYFINSANVSVDPMKSVET
 VKAASLHGSDAIGGVAFETDPAIDPLKNGVDFGQNLQSYSSSDKSPSERVAQAK
 RPDNVALVAYTRDQKELQNSKAPDYVESQDIYVNDLILKQSLQSDLRDLKE
 LGEVLYNQDSDIASSSYKNSQEDTTKQNRGLKHIWRLDTITDSITSLRLDLEK
 ENGLTRFKBASGVPVPPNGDQKQDOYAEKRELETFQDFDELEHYLVVGLS
 YTSRISNTNREPNSDPTPKTYVYTRQDQKQDOYAEKRELETFQDFDELEHYLVVGLS
 DYFSDPSTVDSEPFKESDSAFTGALTYKYAQSGITLFTQISQGFAPSPRELYYT
 YNPQGYTRNRPNPISERSKLSLEYGHNTASTSTALYSDYDTLEQTSNGV
 GLEYVYVNISEATIKIGIEVSNTLIDNRITANPLGWTTFVAVATVGBDQBNPLNSG

NPNAVAVVYDAPNTLMGTSIKVNTYAKKSNHINDGNNGTANOVLEPSATVIDL
TAYFKPDVTVTAAGIMNLITNEKYYLMNDIRGEKEISRLDTQAEFRYNTSVKTEP"
gene
5180. .6094
/locus_tag="VV20277"
CDS
5180. .6094
/locus_tag="VV20277"
/note="COG0583"
/codon_start=1
/transl_table=11
/product="Transcriptional regulator"
/protein_id="AA007242.1"
/db_xref="GI:27358292"
/translation="MILITKGVESVSPITLLEALHIIIDAIRGSPASPAANELRAPS
LSYQIQLEBODIDIMIFDRSGHKANPTDARLILBERGRTILATETCLVNDATITLANG
ELDIISTADIGIHFSLVDALGVNSKTRVLRQEBITLACWESLQSGRADILVCPK
LEVLPHVDKADTIGKMSMTVVAAPNHYNRSGEFDYAEKRIIADTAEGOPAI
TVNIIOKOPRLVTNLQAKVDALVTGIGTLEPTDIATPLIEAGKLEKEGAEQOHLE
IILAKRNQMGKASWCIOYKKQMKLK"
complement(6143. .8821)
/locus_tag="VV20278"
CDS
complement(6143. .8821)
/locus_tag="VV20278"
/note="other COG candidate exists; similar to
acetyltransferase; COG0454"
/codon_start=1
/transl_table=11
/product="Histone acetyltransferase HPA2"
/protein_id="AA007243.1"
/db_xref="GI:27358293"
/translation="MFKLOKLFKPTSIIVAGISORHLSAGNIYMKNLDSGFGDTIIP
VPRYSVAGVLAIVPSVTELPLVPDIAICTRSERVVALFKOLAEIGTEFVILYADT
DYHODQGVADACIARIESNRVYIGRNSLGIILPQKFNASPSPTAKKGIATIS
QSAVCTTILIDANDKEIGSAFVSTGNGLDIDFELDLSTDSKTDAILLYDSIT
DARFTSAAPAAARNRRIILVKGGRSILGGRALNKSADTLDIVDSAIRSGMLAVN
NTHLEPAAVETLHVSPLRGERLAIITNGAPLMAVDTLFERGQLANISSETIEM
SOHLPSMSWNIIDIVDGADKRVATINAILSDCADAVILMSPSAVSDVEFTAO
AIVDAIOSHPRHRRFNILITWMSGOTAKPARMLITKAGIPTYRTPEASVVAFWHLVEY
RRNOKLMEPTPTAPRPLIVSDVNAAKAVSEHLMDEKIAIODTHQGLSILKIFHPLV
PRTIADTSEAVHIAEOIGPVAVKIRSPDIANKSVQGMNLNSQEVASADAIL
DKTLAVPSANIHGLVQGVAKLAGSEERVKVITTKVGPVILLQSGSEWSEADA
LDLPLMLNTLAVYQIVRAIKEGHILQKLPBMDVYGLPAKFLVRIQWVVELPQKE
LIDHPLANGHEFTIIDLADLTLENYLGHGOERLAIIPRPFLVETMLQDGEIILRP
LIDEPESHAEFTINNVSKEDLYKRPSPDGEFHEMLANLTOIDYREMAFVAVSS
GEAKTIGVARAILINPENTAEFAILIRSLDKGKGLKVTLMNKIIDIYCRQKGTAKMTGM
TPTNRGMULTLAKGFSLDIQPEDGADVPLPLNP"
gene
9046. .9621
/locus_tag="VV20279"
CDS
9046. .9621
/locus_tag="VV20279"
/note="COG3266"
/codon_start=1
/transl_table=11
/product="Uncharacterized protein conserved in bacteria"
/protein_id="AA007244.1"
/db_xref="GI:27358294"
/translation="MKKVAITITLSTVLSACASDNTYANTTSHREEFKIEPIPELIM
ABETVQVSVISQGVVMEPQTEKRVVATISPSSETTKAVKLVPSKQODKQORGY
TVQVAVGQAKVDYFASKLPDQGPIMWNTYKIVNGTKWFTVLYGDFATSTAKKAIA
QJPOEFKELOPFKSIDAIKNSHYPTLNKLN"
gene
9788. .10783
/locus_tag="VV20280"
CDS
9788. .10783
/locus_tag="VV20280"
/note="similar to ATP-grasp enzymes; COG1181"
/codon_start=1
/transl_table=11
/product="D-alanine-D-alanine ligase"
/protein_id="AA007245.1"
/db_xref="GI:27358295"
/translation="MNRNVLLCGGSSSEHISLSSEVYLQOGLIENVNVLYKEIK
NMGFPOKERLVYLDIHTKSVKSDENFBSIHIDFVPCIHGFGEGETDIOSLPELAGI
PLTGGCPBASNSFNKITSKLWDAIGINTPTPLITENNESHOISRAEAPKMLTF
VAAARGSSVCGCYTKIEQLSDAIDKAGFSPHQVLEKAVKPRELEVASAYEMNGQIH

ISKPRITAPDGAFYSDKYSAGSHSITTEVAKNLTLEQOLATITOLCSKRVFROWMLR
HLSRIDPFIYSBESEIILNEVNTFPPGNTKLSMEPKKLQHNHKEHETLADICERSL"
gene
11080. .11454
/locus_tag="VV20281"
CDS
11080. .11454
/locus_tag="VV20281"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="AA007246.1"
/db_xref="GI:27358296"
/translation="MAISHYRGNLOSAPNDNTIMOVKIKRVLQGSLSAVKSIDMW
COTASIIDREBSLIDRKQEAATGSSPAENFNCPALINDGEPNAWCFPNGLKIKSK
MAIORITEAYILAKQAEQOK"
gene
11457. .11768
/locus_tag="VV20282"
CDS
11457. .11768
/locus_tag="VV20282"
/note="eIF-1/STU1; COG0023"
/codon_start=1
/transl_table=11
/product="Translation initiation factor 1 related protein"
/protein_id="AA007247.1"
/db_xref="GI:27358297"
/translation="MTLYVSTETGRIKPEEPKASRPKGVIRIOPETGRKGVCV
VTGLDDDAPLKILAAELKKICGGGSVKGDNIEIQDARDKIKTHLEKKGMKVKPAG
G"
gene
12360. .12782
/locus_tag="VV20283"
CDS
12360. .12782
/locus_tag="VV20283"
/note="COG0499"
/codon_start=1
/transl_table=11
/product="S-adenosylhomocysteine hydrolyase"
/protein_id="AA007248.1"
/db_xref="GI:27358298"
/translation="HDFEAFNVLPFRVPSISIAAGAEYFHVQPAIVRMLSGSLPIPNM
Query Match 39.9%; Score 555.6; DB 1; Length 300883;
Best local Similarity 64.9%; Pred. No. 9,6e-97;
Matches 876; Conservative 0; Mismatches 459; Indels 15; Gaps 3;
QY 25 TCGAAGAAATTTGTTATGACGTTATCGCTTGCTTGACGCTTGACGATTTGG 84
DB 118404 TCGATGAAAAAGAGATGACAGTTGGCGATTAACAGTTGGTGGCTTGTCACCG 118345
QY 85 CCGGATGTCGGGGCAGTTGGCAAGCTGGAAGAAAGAAAGATTGGCTTCATCAAG 144
DB 118344 GTTGTACTGCGAAGACTCGCGAGGCTGAAAGAGACTTAAAGTTGGCTTTATCAAA 118285
QY 145 CTCACCGACATGAGCGCGCTGGCGGTGGCGCGCGGAAAGAGCTTCTTCAGAGCAGAGGCG 204
DB 118284 TTGACCGACATGAGCACCGTTGGGAGTCGTTATGAAAAAGATTTTGTAGAGTAGAAGGG 118225
QY 205 CTGTTGTCGAACATGGAAGCGCAGGCAACTGGAAGGTGTGATGATAGGCTCGTAAT 264
DB 118224 CTGTATGATACGTGTAGAAACGAGGCAACTGGAAGGTCTGTCATGCGCGTGAATGAT 118165
QY 265 GGGGAACGTGAGCGCTCGCAACGCTGGCGCGCGGCGCGTATGCGGCGCAGCGTTGGCTTC 324
DB 118164 GGTGAACGTGAGCGCGCGCAACGCTGGGAGGCGCAACCTGGGTGCCAATCGGCATT 118105
QY 325 GGCACCAAGCGCATATCGAGGTGCGCTTCACAGATGGCTTCGAACGCGCAACGCGATTACG 384
DB 118104 GGCACGCAAGCCAAAGTATCAACCGGATTCAGATATGATCTCAATGTATACCGGATCACG 118045
QY 385 GTGTCCATGAATATTCGCATGATGAAGCCGAACATACCGCTGGA---AGCGGTAAA 441
DB 118044 GTGTCCGAAGATGTTGGCAGCAAAATGAAGCCCAATTGTGTCAAGACAGTACGCGTAAA 117985
QY 442 CCGGTGCATCCGATCAAGCAGATTAATCTAAACCGGTGTGGAAGAAATACAAAGCCGAA 501

Db 117984 CCGGTCACCCCAACCAAGCCGATGCGCTTAACCTGCTGTAACAAGTTACCGGACCAA 117925
 QY 502 GCGAAGCCGCTTCAATATGAGGATGACCTTCCCGGCGGATCTCAACAATCAATCCGCT 561
 Db 117924 GCGAAGCCGCTTCAATATGAGGATGACCTTCCCGGCTTCAACAAGTACTGACATCCGCG 117865
 QY 562 TACTGGCTGGCGCTGGCGGTATCAATCTGGCTACTATTTGCGCGCGCA---GACATT 618
 Db 117864 TACTGGCTGGCGCTGGCGGTATCAATCTGGCTACTATTTGCGCGCGCA---GACATT 618
 QY 619 TCGGCGCAATTCGAGCGAGACGCGCTTGTTCGCTGACCGCGCGCGCAATTCGCTGC 678
 Db 117804 AGTGGCGCAATTCGAGCGAGACGCGCTTGTTCGCTGACCGCGCGCGCAATTCGCTGC 117745
 QY 679 ACGCTGAGACCGCGACCAATTTTCGCTTATTCGCTGCGCGAGCGCGTGAACCAAGAGCC 738
 Db 117744 ACCATGGAAGCGGACCAATTCGCTTATTCGCTGCGCGAGCGCGTGAACCAAGAGCC 117685
 QY 739 GTGTTCAAGGCGATAGCGCTGGCGGTGATCAACGATGAAGAACTCTGGAAGAGACCGCG 798
 Db 117684 GTGTTCAAGGCGATAGCGCTGGCGGTGATCAACGATGAAGAACTCTGGAAGAGACCGCG 117625
 QY 799 GAAAGAGCTTCGGCGCGTGAACCAATGAGCGGAGAAATACCCCAACCTATCTGGCG 858
 Db 117624 GAAAGAGCTTCGGCGCGTGAACCAATGAGCGGAGAAATACCCCAACCTATCTGGCG 117565
 QY 859 GTGACCAAGCGCGTGAATTCGGCGCGCGATTCGCTGAGCGCGCAATTAACAAGAACCGC 918
 Db 117564 GTGTTCAAGGCTGATTCGCTGCGCGCGATTCGCTGAGCGCGCAATTAACAAGAACCGC 117505
 QY 919 AAGAAAGCATCAATATGCTGGCGCGCAAAAACAATACGTGCTGCTGAGCTGGAAGTCTG 978
 Db 117504 GCGGAAGCGCGTGAAGCTGCTTTCACGCGAGAGATACGTGGCGCGCGATGCGAAGTATC 117445
 QY 979 GCGGCTAGCATGAACCGCACTTTCGATTAAGAAAAGACATTAACCGCGCGCTACCGGAC 1038
 Db 117444 GCGAATTCATGACTGCGCACTGTTGATTAAGAAAAGCGATTAAGCGCAAGTCCCGAT 117385
 QY 1039 TTCACACCTCTTCTTGGCGCAAGCGCGCGATTCGCTGCTGAGCGAGTCTGCTAT 1098
 Db 117384 TTAAAGTGTCTTCTTGGCTGACCAAGCGCACTTCTGCTGAGCGAGTCTGCTAT 117325
 QY 1099 CTGACCAAGCTGAGCGCGCTGGCGCGATGATCAATGAATTAACCGCGCAACTGCTATCTG 1158
 Db 117324 CTGACCAATATGCGCGCTGGCGCGCAATGCAATCAAGAAACGCGACTGCTATCATG 117265
 QY 1159 GATACCGCGCAAGACGCTTACCGCGCGCGAGATCTATCTGCGCGCGCGCAAGAACTGCTC 1218
 Db 117264 GATACCGCGCAAGACGCTTACCGCGCGCGAGATCTATCTGCGCGCGCGCAAGAACTGCTC 117205
 QY 1219 GCAGAGCGCAAGCGCAAGCGCGAGACTTCCCTGCGC-----GATACCTGATCAAG 1269
 Db 117204 GAAAGAGGAGCATTTGAGCGCGCAAGTTCCTGATCTTTCGCAAGAAAGCGGTTTTCGC 117145
 QY 1270 CCGTGGCAGAAATTTCTTCAATGCAAAAGTCCGCTGATGCAACAAGCGCGCAAGTATAC 1329
 Db 117144 CCACCGCAAAACCACTTCAATGATGATCCGCTATGACGCGCGTACGCCCAATGCTAT 117085
 QY 1330 CTGCGCAAGTTTGGATGATGCTGGAAGGC 1359
 Db 117084 CTGCGCAAGTTTGGATGATGCTGGAAGGC 117055

RESULT 3
 LOCUS AP005348 248850 bp DNA linear BCT 06-JUL-2004
 DEFINITION Vibrio vulnificus YJ016 DNA, chromosome II, complete genome, section 5/8.

ACCESSION AP005348 BAO00038
 VERSION AP005348.1 GI:37201116
 KEYWORDS
 SOURCE Vibrio vulnificus YJ016
 ORGANISM Vibrio vulnificus YJ016

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 SOURCE

Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
 1
 Chen, C.Y., Wu, K.M., Chang, Y.C., Chang, C.H., Tsai, H.C., Liao, T.L., Liu, Y.M., Chen, H.J., Shen, A.B., Li, J.C., Su, T.L., Shao, C.P., Lee, C.T., Hor, L.I. and Tsai, S.F.
 Comparative genome analysis of *Vibrio vulnificus*, a marine pathogen
 Genome Res. 13 (12), 2577-2587 (2003)
 14656965
 2 (bases 1 to 248850)
 Chen, C.Y., Wu, K.M. and Tsai, S.F.
 Direct Submission
 Submitted (29-MAY-2002) Shih Feng Tsai, National Health Research Institutes, Division of Molecular and Genomic Medicine; 128, Yen-Chiu-Yuan Road, Sec 2, Taipei 115, Taiwan
 (E-mail: pctsaio@nhi.org.tw, Tel: 866-2-8146-1041, Fax: 866-2-2789-0484)
 This sequence was determined by the Sequencing Core of the National Yang-Ming University Genome Research Center (YMG);
 http://genome.ym.edu.tw/
 Location/Qualifiers
 1..248850
 /organism="Vibrio vulnificus YJ016"
 /mol_type="genomic DNA"
 /strain="YJ016"
 /db_xref="taxon:196600"
 /chromosome="II"
 383..1786
 /gene="VVA0914"
 383..1786
 /gene="VVA0914"
 /codon_start=1
 /transl_table=1
 /product="anaerobic C4-dicarboxylate transporter"
 /protein_id="BAC96940.1"
 /db_xref="GI:37201117"
 /translation="MNRVYFCTIRENETITNISLRINMLYIEFLFLIMLYISRYEGILGVSGIGLVIEFVFKMPSTSPVVMILIAVVCASILEAAGLKMLOVAERMLRKNKRVTLIAFVYTMFTMGTAIVSIMIIGDVALKNGIRERPMMAASVASOIAITAPISIAVAVVYLAOSDINHETLISILVTPATLFGTILMSLYSMKGKELDEPREVARLKDPMREKINTARTSLDEHLPAARASVLLFTISITIVYIMPEVRIIGEGTPEISMAVVIOMMLCRGGIILTKIDPSRSPGVFKSGMAAIATFGIAMESDTYFYQAMPQFSGIVEMVNNVPMTPALATIVSVVNSQAAVARMPLPVGLGLIEPALVGLMPLVGYFFIPIPSDIATVNFDMVTGTTKIGKWFNHSFMSVGLIGVICACILGLGQMIIIG"
 1858..1891
 /gene="rpt2T08"
 1858..1891
 /note="2 units of 17 bp in-tandem repeat"
 /rpt_type=tandem
 /rpt_unit="ttctattgcgctaac"
 complement(1905..3185)
 /gene="VVA0915"
 complement(1905..3185)
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC96941.1"
 /db_xref="GI:37201118"
 /translation="MMKRLILTLTLQMLQPAVLAAGTNSQYARVWNAQQLADNNALADAIKINSAEVSPPDVAALSKLGIYWOABOFRSIASLEKALSPFLPFEEMVTRMLTQLYLTOGDYKKAIPLNKLKLOYPENESAEIRYLAAGQSYENRETVATKGFQREKEDPTTLISLALAHVQLEOKSVYPLAKQILYKRAQOTLAIENSOYNAVLTNGQESALDMLTLLEQLEINLEATKSLAVLYAAGIYKRAQOTLAIENSOYOPKILKLOAYQWAKEMQOALEMQLRAKQESQYQWEMVLENDLQRYOVLISIAEMDNPKRQPDALQKINALYRLERIDQAVYAKKADALASTKQTMWLKPLSHLKSSBSKLDKAKGQVSEHRLTQQRNVAVKSDSSS"
 complement(3182..3820)
 /gene="VVA0916"
 complement(3182..3820)
 /gene="VVA0916"

LOCUS	302100 bp	DNA	linear	BCT 07-APR-2004
DEFINITION	Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 4/6.			
ACCESSION	AB005087.BA000032			
VERSION	AB005087.1 GI:28809161			
KEYWORDS				
SOURCE	Vibrio parahaemolyticus			
ORGANISM	Vibrio parahaemolyticus			
REFERENCE	Bacteria: Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.			
AUTHORS	1			
TITLE	Makino,H., Iida,T., Sugahara,T., Yamaichi,Y., Park,K.S., Yokoyama,K.,, Uda,T.,			
JOURNAL	Makino,K., Shinagawa,H. and Honda,T.			
MEDLINE	A filamentous phage associated with recent pandemic Vibrio			
REFERENCE	parahaemolyticus O3:K6 strains			
AUTHORS	J. Clin. Microbiol. 38 (6), 2156-2161 (2000)			
	2			
	10834969			
	20295086			
	10834969			
	Makino,K., Oshima,K., Kurokawa,K., Yokoyama,K., Uda,T.,			
	Tagomori,K., Iijima,Y., Najima,M., Nakano,M., Yamashita,A.,			
	Kubota,Y., Kimura,S., Yasunaga,T., Honda,T., Shinagawa,H.,			
	Hattori,M. and Iida,T.			
	Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism			
	distinct from that of V cholerae			
	Lancet 361 (9359), 743-749 (2003)			
	3			
	12620739			
	3 (bases 1 to 302100)			
	Hattori,M., Yamashita,A.; Oshima,K. and Shiba,T.			
	Direct Submission			
	Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for			
	Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara,			
	Kanagawa 228-8555, Japan			
	URL:http://genome.lis.kitasato-u.ac.jp/, Tel:81-42-778-8194,			
	Fax:81-42-778-8193			
COMMENT	This work was done in collaboration with Kozo Makino, Ken Kurokawa,			
	Katsunshi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima,			
	Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga,			
	Takeaki Honda, Hideo Shinagawa, Teruya Iida (Osaka University),			
	Yoshio Iijima (Kobe Institute of Health), and supported by the			
	Research for the Future Program of the Japan Society for the			
	Promotion of Science. This clone was isolated from a patient			
	presenting with acute gastroenteritis.			
	Location/Qualifiers			
FEATURES	1..302100			
source	/organism="Vibrio parahaemolyticus"			
	/mol_type="genomic DNA"			
	/strain="O3:K6"			
	/sub_strain="RIMD 2210633"			
	/db_xref="taxon:670"			
	/chromosome="2"			
	53..1139			
misc_feature	/note="disrupted by frameshift, similar to part of			
	hypothetical protein GB:AA96452.1 (AB004386) percent			
	identity 62 in 165 aa"			
gene	complement (1224..2012)			
	/gene="VPA0861"			
	complement (1224..2012)			
	/gene="VPA0861"			
	/note="similar to GB:AA96483.1 (AE004389) percent			
	identity 73 in 261 aa"			
	/codon_start=1			
	/evidence=not experimental			
	/transl_table=11			
	/product="conserved hypothetical protein"			
	/protein_id="PAC62204.1"			
	/db_xref="GI:28809162"			
	/translation="MKWVIRASALITAGLTACGASPTGRNQIMEDQMSISGAKS			
CDS				

```

gene      2217..2651
CDS       /gene="VPA0862"
          /note="similar to GB:AAF96482.1 (AE004389) percent
          identity 75 in 144 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="BAC62205.1"
          /db_xref="GI:28809163"
          /translation="MSIMKKPISLDILNATSNNTMMHLOIEXTNFTDLSLATMPC
          SPTHQIGMLHGASVYLAETLGSIDANFCVSGNVCGLDINANVTRMMSGYVIGT
          ACPHIGVSTQVWQITETDPRDLVCTSLTIAVKQKAKLA"
gene      2694..2927
CDS       /gene="VPA0863"
          /note="similar to GB:AAF96481.1 (AE004389) percent
          identity 56 in 76 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAC62206.1"
          /db_xref="GI:28809164"
          /translation="WVIEFSLGKIATQREIVIKLTGEAMVTLQAQTDALQLGRGAN
          VVLAHGAECQMSIKLDNEBQLIQLAQEIQIDIQ"
gene      3007..4713
CDS       /gene="VPA0864"
          /note="similar to GB:AAF96480.1 (AE004388) percent
          identity 66 in 565 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="putative ABC transporter periplasmic solute
          binding protein"
          /protein_id="BAC62207.1"
          /db_xref="GI:28809165"
          /translation="MSSPRLRVQFETLFEKPAQNDRETVQLEDITDALFCTRRNARIVL
          NLEBEQWIEHMPAQRGKLSKLIPRANSVDSENARRYLDEBKIGQALDLDMDAA
          RLTOYIGTIGHQROEQVRLPRTPLAMLPQKPKRSQHLARQVSGELTRIDE
          SQLOPLAHATWEALSDTHMRVYLRGVFHNGBPLTTDCVVESILALSGNLFSHIQ
          QVSSQPMWTVDIKLVPRRYLPLALSESQAKVLLPOLASEBDFDRKPIGTGPOYKVN
          DKRLILITAFDGVFGFRPLGGVWVVIDEAYSSWVYPSLSKPKKKQKASTDEVELD
          GCTPLLNKNGTGIADQPRWAEFLSOTLNSHOIYAHVPODKVLEIGVLAQKPKGMD
          LKPTGVGPOQKSTTVAQYKKGHPVYIAKAIKHLKPHSVDEFFIRYDQPPRPE
          EVDVWKAKGATNNRDALAGMLLDYSDIEKTSAGTDFSEMATLVDMQWAGIHADPPA
          RELGKQLVKSQVITPMFHCMLGVNKHDSALQNAKNALGWPDFSVMVKEPIEDNGQ
          AE"
gene      complement(5251..5847)
CDS       /gene="VPA0865"
          /complement(5251..5847)
          /note="similar to GB:AAF96687.1 (AE004407) percent
          identity 71 in 192 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="BAC62208.1"
          /db_xref="GI:28809166"
          /translation="MKSSTLIGLITFAIVFLFSALBESSRAVMPAAIIPIAGILIA
          IWAISIRYQFSMTAYILMTIMLCIHTIGKATPAVRPMPNNLIGSENNNDRAVAF
          AIGLVAIYPAEYLIRKKRNPFTFSCFALFAIMSLAGETIIEWVAELAGDEGIAF
          LGSQGIWDQAQDMLCDITTGALISFLMSAQRFAKPF"
gene      6123..6962
CDS       /gene="VPA0866"
          /gene="VPA0866"
          /gene="VPA0866"

```

```

gene      7095..7715
CDS       /gene="VPA0867"
          /note="similar to GB:AAK22187.1 (AE005694) percent
          identity 30 in 298 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="putative oxidoreductase iron/ascorbate family"
          /protein_id="BAC62209.1"
          /db_xref="GI:28809167"
          /translation="MKLETVDYLDADDAQAFVASIRENGFVGLKXHPPIPOELVESTYK
          NMHEFCSEKNDQFNVETQDGFPPSVSEYAKGHKVKDKIETHVYPMQGFEDLK
          AEILEYERANAFAQELGWBEYAPKQVEKFSIALSEMINNSDKTLRYLHVPMT
          GEEBGLRAAHAEHINLITVLPANBGLQVMSKDGMDIVPCDFGNLIINIGMLQ
          EASGVFSTTHRVINPFGARQEKRSISLPLHPKPDVLSERYTASHYLMERLREL
          GVI"
gene      7095..7715
CDS       /gene="VPA0867"
          /note="similar to GB:AAF96686.1 (AE004407) percent
          identity 63 in 204 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="DnaJ-related protein"
          /protein_id="BAC62210.1"
          /db_xref="GI:28809168"
          /translation="WEDQYGVCDSSQSEIENPLMPMLILKROQSGMKVHTLASKLS
          EKGVINQDPSPKDLFRANFLINALLQLOQTLPDLSWLOVEANDIMLSALESRN
          SYDPNDPLROYLIDWSHYEANDDEVRLNEFWTLPYRQVVGASHNIDKAKALKLF
          LAGDYSRTIEIRQMRKALRMHPDRENGDAERFVLCBAMVNLKNG"
gene      complement(7844..9499)
CDS       /gene="VPA0868"
          /complement(7844..9499)
          /note="similar to GB:AAF96684.1 (AE004407) percent
          identity 78 in 548 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="BAC62211.1"
          /db_xref="GI:28809169"
          /translation="MDQLTATLTKIKONRYAYQOIKGOYDPTFTLFDHVGDPYA
          SASRRATRAWSLTLEMLKDESPAFOBAARDPIARSQEPKQKNTSIALNGOTV
          DSTAYLFTBEGIELEFRVNLPAHGSVAGKANNLITLHPLEFTRATIERLDEAM
          VHCQVVDQSLAREQLBAHNLVAVNANGSVLPRIAGCNDLPKKAVERTAESIQVT
          LHAENKGVYTGIGIKGTLLVGGGFHOKSVTLNAIESIYDHIGDGEYLVTDQKA
          NKIRAEDEKCVHNLMSVYNLPMGKTDAPSTQDASGSTQAAMLQESIEAGATSL
          LIDETSATNPMIRBERMOALVAKGDEPITPLVDRIGLRDLDSITIIWVGSGGYL
          DVANVVICMHDVQADVTEKAKQVIAQHPQGHNSSESLQTFRRPALNRVALNMLIT
          DKFRVSAKGDSLFPKGFPTDLSALBQIESADEVNAIGMLFOLAQLPGKCNPAKE
          IEEMLSGEMHSLPQSGDLAKPRTLDVMAALNRKSKQPKPSH"
gene      complement(9571..11481)
CDS       /gene="VPA0869"

```

Query Match 37.1%; Score 516.2; DB 1; Length 302100;

Best Local Similarity 64.4%; Pred. No. 3.6e-89;

Matches 827; Conservative 0; Mismatches 443; Indels 15; Gaps 3;

```

QY 101 TTGGCAAGCTGGAAGAAGATTGAATTGGGCTTCATCAAGCTCACCGCATGGCGC 160
DB 217725 TGGGTGACACAGAAATTGAAGACCTGMACTTGTTTATTAAAGTTAACCGATATGGCGC 217666
QY 161 CGCTGGCGGTGGCCGCCGAAAAGGCTTCTTGAGAGACGAGGCGCTGTTGCGTCAACTGG 220
DB 217665 CGTTAGCGGTTGCTTACGAGAAAGGCTTTTGAAGACGAAAGGCGTATACGTTAGCGTGG 217606
QY 221 AAGCGGAGGCGCAACTGGAAGGTTGCGTGAATAGAGGTCGTGATGCCAATCGACGCGT 280
DB 217605 AAGCGGAGGCGCAACTGGAAGGTTGCGTGAATAGAGGTCGTGATGCCAATCGACGCGT 217546
QY 281 CGCACATGCTGGCGCGCGGCGCTTAGCGGCGCAGGCGTTCGCGACCAAGCGCGATA 340

```


Db 217545 CGCATATGCTGGACGAGGACCCATTTGGGCGCAACGATTTGTCGGACCAAAAACGAG 217486
 Qy 341 TCGAGGTGCGCTTCAAGCATGGGCTTCAACGCGCAACGCGATTACGGTGTCCAAATGAATCT 400
 Db 217485 TGAATACCGGCTTCAAGTATGATTTGAACGGCAACGCAATCACTGTTTCCAAATGACAGT 217426
 Qy 401 GGCATCATGATGAAGCCGCAACATAC--GCTGGAAGCCGCTAAACCGGTGATCCGATCA 457
 Db 217425 GGGAAACAGATGAAGCCCAACATACCAAAAGACGAGATGGCAAACTGTATCATCTCAATCA 217366
 Qy 458 AGCAGATTAATCTAAACCGGCTGTGCAAAATACAAAGCGAAGGCAAGCCGTTCAATA 517
 Db 217365 AACCGATTTGTTGAAGCCCGCTGTGACAGTTATCGCGCAAGGCGAAGTCAATTAATA 217306
 Qy 518 TGGCGATGACCTTCCCGGCGGATCTTCAACATCAAACTGCTTACTGCTGGCGGCTG 577
 Db 217305 TGGGCAATGGTGTCTTCTGTTTCACTCACAATTAACGATTGCGTTATTTGGCTTGTGCTG 217246
 Qy 578 GCGGTATCAATCTCGGTACTATTTCGCG--GCGGACAGACATTTCCGCGCAATTCGCG 634
 Db 217245 GTGGGATTCATCCAGGTTTATATGACACAGACTCAGGCGAATAACAGTGTCAATTTGATG 217186
 Qy 635 CAGACGCTTGTGTGGGTGACCCCGCGCGCAAAATCCGCTCAGCGTGAAGCCGCGCA 694
 Db 217185 CTGACGTTCTTTGAGTGTACACCGCACCAAAATGCGACCAATGAAAGCGGCA 217126
 Qy 695 CCATTTTCGTTATTTGCGTGGCGGAGCCGTGAAACAGACGCGTGTCAAGGCGCATAG 754
 Db 217125 CATTCAAAGGATCTGGTGTGGGCAACATGAAACGACGACGCGATTTTCAAGGCGATTG 217066
 Qy 755 GCGTCCCGTGTATCACCAGTAAAGAACTGTGAAGAGACAGCGCGGAAAAAGTCTTCGCG 814
 Db 217065 GTACGCGACAGTCGACCGATTAAGAAATCTGAAAGAACCAAGAAAGTGTGTTG 217006
 Qy 815 TGAACAAACATGGGCGGAAATAATCCCAACACTATCTGGGCGTGAACCAAGCGCTGA 874
 Db 217005 TGTCCAAAGATGGGCGGAAATAATCCCAACACTATCTGGTGTGTAAGGCGCTGA 216946
 Qy 875 TTGCGGCGCGATCTGGCTGAGCGCGCAATTAACAAAGCCGCAAGGAGCCATCGAAA 934
 Db 216945 TTGACGCGGACATTTGGTGTGATGAAACAGATGCAACCGCCAGAAAGCGGTGAAGA 216886
 Qy 935 TGTGCGCGCAAAACATATCGTGGTGTGACGTTGAAGTGTGGCGGCTAGACATGAACG 994
 Db 216885 TGTGCGCAAGACGAGTATGTGGAGCGCATGCCATGTTATTCGCAACTCATGACAG 216826
 Qy 995 GCACTTTCCGAATACGAAAAAGATTAACGCGCGCTACCGGACTTCAACCTTTCTTC 1054
 Db 216825 GCACTTTGAATACGAAAAAGGAGCAACAAAGACGTAACCACTTCACGTTCTTCC 216766
 Qy 1055 GCCACGCGCGCAGCTATCCGCTTACAGAGTGCAGTCTGTATCTGACCCAGTGAAGC 1114
 Db 216765 GTATTAACGCAAGTACCGCTACTACGCAATGCGATTTGTTATCTGACCCCAATGCGCC 216706
 Qy 1115 GCTGGGCGATGATCAATGAATTAACCGGACAACTGTTATCTGATACCGCAAGAACG 1174
 Db 216705 GCTGGGAGCAAAATCTCAAGGAGAAATCCGATGATTTGTAATGATGTGGCAAGAGG 216646
 Qy 1175 TCTACCGCGCGGACATCTATCTGCGCGCGGCAAGAACTGTGCGCAAGAGCGCA 1234
 Db 216645 TTATACGTCAGATCTTACCAACAGCACTCAAGGTTAATGAAGATGCTGTGTTGA 216586
 Qy 1235 AGGCGGAGACTTCC-----TGCGGATACCTGATCAAGCGGTGCGAGATTTCT 1285
 Db 216585 GCAAAAAAGACTTCCAGATTTCAAGTGTGAAGATGTTTTCGTCGCCCAACAAACACT 216526
 Qy 1286 TCAATGCAAAAGTACCGTTGATGCAACAAAGCCCAACGATTAACCTGCGCAAGTTTCCA 1345
 Db 216525 TTATTCGCAACATCGTTATGAGGTGCGAGCGCAATAATATCTGGAAGAAGTTCTCTA 216466
 Qy 1346 TAGCTCTGAAGGCAAGCAAAACGT 1370
 Db 216465 TCGGGCTTAAAGAAAAAGCAAAAGT 216441

RESULT 5
 BX294133
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

BX294133 311050 bp DNA linear BCT 11-JUL-2003
 Pirellula sp. strain 1 complete genome, segment 1/24.
 BX294133 BX119912
 BX294133.1 GI:32442811
 complete genome.
 Rhodopirellula baltica SH 1
 Rhodopirellula baltica SH 1
 Bacteria; Planctomycetes; Planctomycetaceae; Planctomycetales;
 Planctomycetaceae; Rhodopirellula.
 1 (bases 1 to 311050)
 Gloeckner, F.O., Kube, M., Bauer, M., Teeling, H., Lombardot, T.,
 Ludwig, W., Gade, D., Beck, A., Borzym, K., Heilmann, K., Rabus, R.,
 Schlesener, H., Aumann, R. and Reinhardt, R.
 Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1
 Proc. Natl. Acad. Sci. U.S.A. 100 (14), 8298-8303 (2003)
 22735913
 12835416
 2 (bases 1 to 311050)
 Kube, M., Borzym, K., Heilmann, K., Klages, S., Margardt, I.,
 Lehnack, S., Beck, A., Pawlik, R., Reinhardt, R., Gloeckner, F.O.,
 Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Rabus, R.,
 Schlesener, H. and Aumann, R.
 Direct Submission
 Submitted (21-JAN-2003) Max Planck Institute for Molecular
 Genetics, Prosektasse 73, D-14195 Berlin, Germany Max
 Planck Institute for Marine Microbiology Celsiusstrasse 1, D-28359
 Bremen, Germany
 This project was carried out by
 *Max Planck Institute for Molecular Genetics, Berlin, Germany; *Max
 Planck Institute for Marine Microbiology, Bremen, Germany; in the
 framework of the RBGX-project, <http://www.regx.de> -----
 Genome Center
 Center: Max Planck Institute for Molecular Genetics
 Center code: MPIWG
 ----- Summary Statistics
 Sequencing vector: pUC19; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 7142841 bases at least Q40
 Consensus quality: 7145138 bases at least Q30
 Consensus quality: 7145484 bases at least Q20
 Quality coverage: 8.03

This sequence was finished as follows unless otherwise noted: all
 regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid sequence; assembly was additionally confirmed by long
 range PCR and cosmid end sequences.

See <http://www.micro-genomes.mpg.de/pirellula/> for more information
 including minimal tiling path from a set of 220 cosmids out of
 908. See the misc_feature tag below for the boundaries of the MTP
 cosmids. ----- Annotation
 Center: Max Planck Institute for Marine Microbiology
 Celsiusstrasse 1, D-28359 Bremen, Germany.
 Celisuserbase 1, D-28359 Bremen, Germany.

Center Code: MPIWG
 Email: fog@mpi-bremen.de
 Phone: +49 (0)421 2028 938 Fax: +49 (0)421 2028 560

Three different programs (Glimmer, Critica and Orpheus) were used
 for ORF-prediction. A nonredundant list of ORFs was generated by
 suitable parsing of the results.
 Automated annotation was done with the software package Pedant Pro
 (<http://www.biomed.de>). All ORF predictions and annotations were
 manually corrected by considering all results of the different
 tools applied. See <http://www.regx.de> for more information and

access to supplementary information.

FEATURES
source

Location/Qualifiers
1. .311050
/organism="Rhodopirellula baltica SH 1"
/mol_type="genomic DNA"
/strain="1"

gene

/db_xref="taxon:243090"
complement (3. .182)
/locus_tag="RB2"
complement (3. .182)
/locus_tag="RB2"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein-signal peptide and transmembrane prediction"
/protein_id="CAD71330.1"
/db_xref="GI:32442812"
/translation="MSRLPFSALFAVSVFVTCGGGGEVIEDTRSPAEVQEGED YEKMBETMDSATNS"

gene

39. .260
/locus_tag="RB3"
39. .260
/locus_tag="RB3"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein-signal peptide and transmembrane prediction"
/protein_id="CAD71331.1"
/db_xref="GI:32442813"
/translation="MLFVIFLLHFGMRAGVFDHVAASAATNSDRVAERKQGEK OSRHYYCSMDANECKVRELILRISP"

gene

complement (399. .1514)
/locus_tag="RB5"
complement (399. .1514)
/locus_tag="RB5"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein-signal peptide and transmembrane prediction"
/protein_id="CAD71332.1"
/db_xref="GI:32442814"
/translation="MTKRLPTRORHVLVAAGLITPLAASAKAGFGMEWETGSPY YEDDAWYSEWFDGNDYPTBPAICRMNENYTDREALTSDDNDIDITWKDGYT AITDSDDTWYIDHGYSDGSDNNDGVEYTSNYDNDGVDAFSEYDVGDG YLDDATVYISFNSABQNDQKTAQKADQNASAKSQSTGTITRAKKTATTNV VYLQNSQSEKSMIALDGPADQDLDMFQMLNDLTVGSGPKAGDTSLRSIEHNGQ MOVERSGREYSGTISTGTVDVRRGOOROLAKITDDGKMLVDLGRKQDLDTNTEGS OYTINGPAQVDRWMLATKTYTLDSGDTIEISMAQK"

gene

1548. .1706
/locus_tag="RB7"
1548. .1706
/locus_tag="RB7"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein"
/protein_id="CAD71333.1"
/db_xref="GI:32442815"
/translation="MRSRSQSVVOETNGRKVDKADKIGARVAQPNRFEGKLSGVAAG RGGKQKVO"

gene

complement (1733. .5263)
/locus_tag="RB8"
complement (1733. .5263)
/locus_tag="RB8"
/function="DNA synthesis and replication; transcription"
/note="best DB hit: BLAST: pir:DB3545; probable helicase PA0799 [imported] - Pseudomonas; E=7e-95 embi:CA67095.1; (X98455) SNF [Bacillus cereus]; E=8e-94 pir:CT2027; sw:snf family helicase 2 - Chlamydomophila pneumoniae; E=4e-87 COG: PA0799; COG0553 Superfamily II DNA/RNA helicases; SNF2 family; E=7e-96 PFAM: PF02803; Thiolase, C-terminal domain; E=0.11 PF00176; SNF2 and others N-terminal domain; E=3.9e-64 PF00271; Helicase conserved C-terminal doma;

CDS

E=1.2e-24"

/codon_start=1
/transl_table=11
/product="helicase"
/protein_id="CAD71334.1"
/db_xref="GI:32442816"

/translation="MDWIAEDQSLGILMEAVRVLVDSHDGFSRAAMLTVTKPT LTQDNGVRFPQVVGNNRKKQTVLALVLPQVDEDDPGGVELSSECCGCAFLBDN SCPTHLLAAMWLOEQIARRNGGQLEIFAGGKAKAVALAKQFVBLLEARRANSRAP AEDAPTRLOMVRVYMAQOTAPLTTFPEQKQGGKMTGTRVRSFGDLNDONLT DPLDSKISLSTVPTGHVDNVAIFEAALPMAGHENVAMDONALPVFVCSQLLT LQPEVBOSSKKPKASKSQSDGSKSKAAKTRTSSGSGSKGDELIGNTAYATMT PESLDSIEASADDELIPDLIPDEKVVQOPRLSDGIDIDIQKCEIVVGLSPV EPAITLADDEDRIVLCTADPAAGITIOFLKRGYRDLTHATASRLAALAGSIDS LIRBLPQLAGPIVAPCELVQMRPRGAGLKVSLRYESRLDSLLPGAPELVP CLAEQPIRLQDLKAELESARKIVQRKLSLSHSTHEWALTDHAAIDLGLRHD ADESPRIYVPEGETFVRGETTPKSLRKIDSDPMQMTGTVDGVDVLODLM AYREDROLVVRDDRPAKISEAFRRKLTQGLTVSEKGLDLAGAVALVDVIGDD VTLRTTAHQHAIERLDSIQMSPPRPSGLDNTLTDYQLDGRMLARLSRMVGVYLA DMGLKTVQTLIVERGPVGPALVVAPTSIGENWVAEKFETALTKLRYDCDLO EVIKTAGPNDLLIVSYOLQORDAKRFAARGHLLVDAQFKNQSTQSAIRITQA DWRILSGTPIENHGLBELMISLFRLLSPGLSGMORPSRFAPIERHSPADRESLAR LVRPVLRRTKDKVLEKLPPEPTEITLRAELSAPEKLYEDARLALAEITAGGAPSH EGRRRLOTLISWTLRLROLAHPISLVEPMKTSKLOFLSLVEELREGDHRALVPSO PYKHLISVRAALDESGISYQYLDGATPSHERORVDAQONGGDLFLISIKRGTCIN LTRADYVLAHDPWMPVAVEDQATDRAHRTQGERAVTVRLVAERTIEQIOLDAHDKR ELVAGILDTDAALQTDLLIDLLIKAEWG"

complement (5458. .5850)
/locus_tag="RB13"
complement (5458. .5850)
/locus_tag="RB13"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein"
/protein_id="CAD71335.1"
/db_xref="GI:32442817"
/translation="MHLPHGNRFPKDRKRYRETSFTLLRTLATETVRLPVAHPF PRSFELMETPSSDPQSLSKDCPKIPSGTIPENKVCSPSVIPKELPHTTCEOPPLSNR AGNNRVLDKRGVWGFGTRVPLRLSP"

gene

5836. .7386
/locus_tag="RB14"
5836. .7386
/locus_tag="RB14"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein"
/protein_id="CAD71336.1"
/db_xref="GI:32442818"
/translation="WGMHPVYDPDANFPFESTREWEDAYLCIEGVPMQSRYSNGR FTLRGVDTSAKLFLEPSTIGRQUSLCSLRCDRSHRLVEALAEALHSNEVGLD VWORGVLVSRFDFDLRTGRFLDPAQCAAPDSCVDALAEALAEALHSNEVGLD FSTQSVAYRSRERLSTMLAVGVLPPIPAIEVTDMDQSGASNSIDEQATITRAFNA AAVRISMGELIETDSGRPNVEPTDEALTCQAQGVLEIGAPVLYDKGLPDLTLDD DEDKLSMTSIEVETKTVQFRRGVVLMNAASINAVGIPGPDQWEPISGIILOTIR RCDPSTPMSLSDPCGCEYLAARDGSGISPIPHADLALSGLAGISIGUNFRGPRDD THPRSAVEFGQWIDRWGTLNAPLVQSLVAGAPGDAARLPIQTRPLSPETNGRPPDA WAKAQIEFVQPLIETLLSKQIVHAVVWDGSDQHPHMTPHAGVIDAKGRPRMLAYLT KIREMLI"

gene

complement (5458. .5850)
/locus_tag="RB13"
complement (5458. .5850)
/locus_tag="RB13"
/locus_start=1
/locus_end=11

CDS

/product="hypothetical protein"
/protein_id="CAD71336.1"
/db_xref="GI:32442818"
/translation="WGMHPVYDPDANFPFESTREWEDAYLCIEGVPMQSRYSNGR FTLRGVDTSAKLFLEPSTIGRQUSLCSLRCDRSHRLVEALAEALHSNEVGLD VWORGVLVSRFDFDLRTGRFLDPAQCAAPDSCVDALAEALAEALHSNEVGLD FSTQSVAYRSRERLSTMLAVGVLPPIPAIEVTDMDQSGASNSIDEQATITRAFNA AAVRISMGELIETDSGRPNVEPTDEALTCQAQGVLEIGAPVLYDKGLPDLTLDD DEDKLSMTSIEVETKTVQFRRGVVLMNAASINAVGIPGPDQWEPISGIILOTIR RCDPSTPMSLSDPCGCEYLAARDGSGISPIPHADLALSGLAGISIGUNFRGPRDD THPRSAVEFGQWIDRWGTLNAPLVQSLVAGAPGDAARLPIQTRPLSPETNGRPPDA WAKAQIEFVQPLIETLLSKQIVHAVVWDGSDQHPHMTPHAGVIDAKGRPRMLAYLT KIREMLI"

Query Match 31.0%; Score 431; DB 1; Length 311050;
Best Local Similarity 62.6%; Pred. No. 8.9e-73;
Matches 747; Conservative 0; Mismatches 425; Indels 21; Gaps 4;

QY	79	ATTGCGCCGATGTCGGGACAGTGGCAAGCGAAGAAAGAAAGATTGGAATTCGGCTTC	138
DB	184764	AATCAACGAGAGTCGTCGCAATGTGGACCTTGGAGAAATCGATTGACATTCGGCTTC	184823
QY	139	ATCAAGCTCAGCAGATGCGCGCGTGGCGGTGCGCGCGAAGAAAGCTTCTTCAAGAC	198
DB	184824	ATCAACGACGACGACGCTCTTGGTGAATCGCAAGAAAGGCGTACTTCCAGCAC	184883
QY	199	GAGGCGCTTGTGTCAGCTGGAAGCGCAAGCTGGAAGCTGTATGATGAGGTC	258

Db 184884 GAAGACTGACGTCATTGGAAACCCAGTCGAACTGGAAGATCTTCTGACAAAGTC 184943
Qy 259 GTGATGCGCACTGACGCGCTGCGACATGCTGCGCGCGCGCTTAGCGGCAAGCTT 318
Db 184944 ATCAATGGCACTCGCGCGGCAACATGTTGGCCGCGCAACCGATCGGTGCAACGATC 185003
Qy 319 GCGTTCGGCAACCAAGCGGATATGAGAGTCCGTTGAGCATGGGCTTCAACGGCAACGG 378
Db 185004 GGTGTGGAAACCAAGTCGCCAATCGTACCGCTTACAGCTTGAACAACGGCAATGGC 185063
Qy 379 ATTACGGTGCATGAAATCTGGCATCAGATGAAGCGCAACATACC---GCTGGAAGGC 435
Db 185064 ATCAACGTCGACGACGAAATTTGGGCAAGATGACGAGAAACGATCTTGAATGGAAGAC 185123
Qy 436 GGTAAACCGGTGATCCGATCAAGGACGATTAATCTAAACCGGTGTCGAAATAATAC--- 492
Db 185124 CCGACACCGAAACATCCCATCAGCGCGCGCAGCTGAACCGATTTGTCGAGATTAACCTG 185183
Qy 493 AAAGCGAAGGCAACCGCTTCAATATGCGATGACCTTCCCGCGGATCTCACACATC 552
Db 185184 CAAGACGCGCGCAACCATTCGCGATGGCATGGTGTTCCTGTTACACTCAATTAAC 185243
Qy 553 AAATCGGTACTGCGCGTGGCGGCGGTGATCAATCTGCTACTATTGCGCGCGCAG 612
Db 185244 GAGATTCGATATTGGCTGGCGCGGTGACGAGATTCATCCCGCATGTACACCGAATC---T 185300
Qy 613 GACATTTCCGCGCAATTCGCGGACGACGCTTGTGTGCGTGAACCCGCGCGCAATG 672
Db 185301 GACATCAAAAGCTTCAACGACGACGATGCAAAATGTGTCAGCGGTGACACCGCTCAATG 185360
Qy 673 CCGTCAACGCTGAAAGCGGACCATTTTGCTTATTCGCTGCGGACCGGTGAAACGAG 732
Db 185361 CCTCAGAACTGGAAGGCGACATGTCMAAGTTACTGCTGTGCGAAACGCTGGAACMA 185420
Qy 733 CAGCGCGTGTCAAGGCGATAGGCGTGGCGGTGATCAACGATGAAGAACTCTGGAAGAC 792
Db 185421 AAAGCGTGTGTGCTGCGCATCGGTGTTCAGTCCACGAACTACGACATCTGGAAGAC 185480
Qy 793 ACCGCGGAAAAAGTCTTCCGCGGTGACCAAACTAGGCGGAGAAATCCCAACACTAT 852
Db 185481 AACCTGAAAAAGTCTTCCGCGGTGACCGGAAATGGCGGAGAAATCCCAACACTAC 185540
Qy 853 CTGCGGTGACCAAGCGCTGATTCGGGCGCGCATCTGCGTGAACGCGCAATTAACAG 912
Db 185541 TTGGCGGTGATCAAGGCTGATCCGCTGCGCAAAATGCTGATGTCACCGATGACTCG 185600
Qy 913 -----AACCGCAAGAAAGCCATGGAATGCTGGCGCAAAAACATACGTGCT 960
Db 185601 GGCATAATGCTCAATCGCGCAAGAGCGGTGAAATCTCAAGCGCAAGATTAATGCTGCT 185660
Qy 961 GCTGACGTGGAAGTGTGCGCGGTAGACATGACGCGCACTTTCGAATTCGAAAAAGACAT 1020
Db 185661 GCGGAGAAAGGTGATTTGGCAACTGATATGAGGACGTTTGTCTTCCATTCACCGAC 185720
Qy 1021 AAACGCGCGCTACCGACTTCAACACCTTCTTTCGCCACGCGCGCACTATCCGCTTAC 1080
Db 185721 GTTTCGAGATGCGCGACTTCAACGTTCTTCAAAACAGAGCGACTACCGCACTAC 185780
Qy 1081 AGCATGTGCACTGTGTATCTGACCCAGCTGAGGCGCTGGGCAATGATCAATGAATTCAA 1140
Db 185781 AGGATGCGCAATTTGTCTTCTCAACAGATCGACGCTGGGGAACAAATCAAGATTCAAA 185840
Qy 1141 CCGGACAACGTGATCTGATACCGCAAGAGTCTACAGCGCGCGCACTATCTATCTGCGC 1200
Db 185841 CCGGCAAGTTGTATCGCGCAAGGCGCAAGAGATCTACAGCGCTGAGATCTATCGCCAA 185900
Qy 1201 GCGGCAAGAACTGTGTGCAAGAGGCAAGGCGCAAGGCGCAAGACTTCCCTGCG 1253
Db 185901 GCGGCGAGGCTTCTAATCAAGCAAGGCAACTGATCAATGAATATCCCGGC 185953

RESULT 6
AB017282_06/c

WPCOMMENT
Sequence split into 33 fragments LOCUS AB017282 Accession AB017282
Fragment Name Begin End
AB017282_00 1 11000
AB017282_01 100001 210000
AB017282_02 200001 310000
AB017282_03 300001 410000
AB017282_04 400001 510000
AB017282_05 500001 610000
AB017282_06 600001 710000
AB017282_07 700001 810000
AB017282_08 800001 910000
AB017282_09 900001 1010000
AB017282_10 1000001 1110000
AB017282_11 1100001 1210000
AB017282_12 1200001 1310000
AB017282_13 1300001 1410000
AB017282_14 1400001 1510000
AB017282_15 1500001 1610000
AB017282_16 1600001 1710000
AB017282_17 1700001 1810000
AB017282_18 1800001 1910000
AB017282_19 1900001 2010000
AB017282_20 2000001 2110000
AB017282_21 2100001 2210000
AB017282_22 2200001 2310000
AB017282_23 2300001 2410000
AB017282_24 2400001 2510000
AB017282_25 2500001 2610000
AB017282_26 2600001 2710000
AB017282_27 2700001 2810000
AB017282_28 2800001 2910000
AB017282_29 2900001 3010000
AB017282_30 3000001 3110000
AB017282_31 3100001 3210000
AB017282_32 3200001 3304554
Continuation (7 of 33) of AB017282 from base 600001 (AB017282 Methylococcus capsulatus str
Query Match 12.3%; Score 171.2; DB 1; Length 110000;
Best Local Similarity 53.3%; Pred. No. 8.4e-23;
Matches 542; Conservative 0; Mismatches 408; Indels 66; Gaps 6;
Qy 125 TGAATTTGGCTTTCATCAAGCTGACCGACATGCGCGCTGGCGGTGCGCGGAAAAAG 184
Db 18318 TGAACGTGGTTTCATTCCTCCTCAGGACTGCGCTCGCTGTCATGCGCTGGAAG 18259
Qy 185 GCTTCTTCAAGGACGAGGCGCTTGTGTGCACTGGAAGCGGCAACTGGAAGGTGG 244
Db 18258 GTTACTTCCGCGCACAGCGGCTGGAAGTCACTGTGCGCGCACGCGTCTTGGCGGCCA 18199
Qy 245 TGATGATAGGGTGTGGAATGCGGAATGCGGAATGGAAGGCTGCGCATGCTGGCGCGCGCT 304
Db 18198 TCGCGACAAAGGTGCGGACCGGCTGTGGAAGCGGCCACATGCTGCGCGCATGCGCA 18139
Qy 305 TAGCGGCGAGCGTTGGCTT---CGGACCAAGGCCGATATCGAGTGGCGCTTGACATGG 361
Db 18138 TCGCGACACCTTCGCGCTGGGGGGGTACGGAAGGCTACATTTGGCGCATGCTGG 18079
Qy 362 GCTTCAACGCAACGCAATTAAGGTGTCGATGAAATTTGGCATTCAGTAAGCGCAACA 421
Db 18078 ATCTGATGGGAACGCGGTGACGTTGTCACACGCTGTAC-----GGCGCTCA 18028
Qy 422 TACCGTGAAGGCGGTAAACCGGTGATCCGATCCGATCAAGGAGATTAATCTAAACCGGTG 481
Db 18027 CCGCGCGGCGCGGCTGAGCTTCCCGCAGGCGCGGTTTGGCGCGCTGCTCGGCGTT 17968
Qy 482 TCGAAAAATGCAAGCGCAAGGCAAGCGCTTCAATATGCGATGACCTTCCGCGCGAT 541
Db 17967 TCGTGAACAGGACGCGGAGGTCTCTGCGTTTGCACAGGATTCCTCTTCTC-----T 17914
Qy 542 CTCAACAATCAAACTGCTTACTGCTGCGCGCTGGCGGTGATCAATCTGCGCTACTATT 601
Db 17913 CGCACAATCACTGCTGCTGCTTGTGCTGCGCGCGCGGCGGTGATCC----- 17865

QY 602 CGCGCGCGGAGACATTTCCGGCCAAATCGGCGCAGACGCGCTTGTGTGCGTGACCCCGC 661
Db 17864 -----CGACCGCACGTCGGCATGACGCTCATTTCCCG 17833
QY 662 CGCGCGGAAATGCGGTCACGCTTGAAGCGCGGACACATTTTCGCTATATGCGTGGCGAG 721
Db 17832 CGCGCGAGATGCCAGCTTGCCTGCGCGCGGCGGAGATTTCCGCTACTGCGCGGCGTGAAC 17773
QY 722 CGTGAACGACGACGCGCTGTCTTGAAGCGCATGAGCGCTGCGGTGATCACCGATGAAGAAC 781
Db 17772 CCGGATATCCCGCTGCGCGCGCCAGGCGGATCGGCGGTGATCATGATCACCGGTGATGAAA 17713
QY 782 TCTGGAAGACACGCGCGGAAAAAGTCTTGGCGCTGACCAACATAGGCGGAGAAATATAC 841
Db 17712 TCTGGAACACAGTCCCGAAAAAGTCTGCGCTGACCAAGGCTGAGCGCGCAACATATC 17653
QY 842 CCAACACCTATCTGCGCGTGAACCAAGCGCTGATTTGGCGCGCGATCTGGCTGAGCGCG 901
Db 17652 CCAATACCTTGGAGAGCCCTCTCATATGCGCTGCTGGAAGCGCGCATGCT-----CG 17599
QY 902 ACAATACAGAAACCGCAAGAAAGCCATGCAATGCTGGCGCAAAAAACAATACGTGCGTG 961
Db 17598 ATCGGCGCGCAAAACCGCGCGAAGCGCGCGCTCTCTCGGAATCCGACTATGTGCGCG 17539
QY 962 CTGACCTGGAAGCTGCGCGGCTAGCATGAACGCGCACTTTGCAATACGAAAAAGACGATA 1021
Db 17538 TCGCGCGGAGATCTTCAGATGTCATGACGAGTTCGTTCAATATGCGCGCGGCGAGAC 17479
QY 1022 AACGCG---CGTACCGGAGCTTCAACACCTTTCTTGGCAAGCGCGCGAGCTATCGTCT 1078
Db 17478 CGGCGCTGTCCTCCCGGACTTCAAGTGTTCACCGCTATGCGCGGACCTTCCCTGGC 17419
QY 1079 ACAGCAGTCACTGTGTATCTGACCCAGCTGAGGCGCTGGGCGCATGATCAATGAA 1134
Db 17418 GTTCCACCGCCCTGTGTCTTGAACCAAGATGCTGCGTGGGCGAGATCGCGAA 17363

RESULT 7
AE008293 10676 bp DNA linear BCT 28-MAY-2004
LOCUS Agrobacterium tumefaciens str. C58 187 of the complete sequence.
DEFINITION
ACCESSION AE008293.1
VERSION AE008293.1 GI:15159385
KEYWORDS
SOURCE Agrobacterium tumefaciens str. C58
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE 1 (bases 1 to 10676)
Hinkle,G., Slater,S.C. and Goodner,B.
Complete Genome Sequence of Agrobacterium tumefaciens C58 (Rhizobium radiobacter C58), the Causative Agent of Crown Gall Disease in Plants
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 10676)
Hinkle,G., Slater,S.C. and Goodner,B.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Bioinformatics, Cereon Genomics, 45 Sidney Street, Cambridge, MA 02139, USA
COMMENT Approximately 800 bp of telomeric sequence missing from the left end of the chromosome and 200 bp missing from the right end.
FEATURES
source
1..10676
/organism="Agrobacterium tumefaciens str. C58"
/mol_type="genomic DNA"
/strain="C58"
/isolate="Cereon"
/db_xref="taxon:176299"
/complement(93..1352)
/gene="AGR_L_1874"
/complement(93..1352)
CDS

gene
CDS
/gene="AGR_L_1874"
/note="HYPOTHEITICAL 40.1 KDA PROTEIN IN SIFU-KIP1 INTERGENIC REGION"
/codon_start=1
/transl_table=11
/product="AGR_L_1874p"
/protein_id="AAK89510.1"
/db_xref="GI:15159386"
/translation="MERKSMEOKKLTPTPADLMSKRSALGALFIMATSAIPGFEITQTATPTDGAFAFGILASLIDFVVDLNTWRIVTLTMRASDIANAIIPAGYFLAILIIGLFFNIGIGTGLJNALFGDLPKIGASIAISIAIVSKAGALVAFRIIPAGIMIVLTLYVAFVSAPVGDAROFPLEDTINPATITTIIGVGVGTYTSGAHRLDGMVGIENLPVNRBALGIVTGMRYVLFLAILGVVAGVVIDTSGAANPAAQPSAAGDLDFRLFGIIFMAAITSVIGAAVTSVSPLPVFKOMSEBAARMATVPIAVSLVCLITTPPAALVFGVGLNGILPLIGLSIFLPAWKRBDLNGHRYSRILLGIVLTICALTWGYSKAGTIFGLGL"
1517..2266
/gene="AGR_L_1876"
1517..2266
/gene="AGR_L_1876"
/note="(ALJ3213) putative GntR-family transcriptional regulator."
/codon_start=1
/transl_table=11
/product="AGR_L_1876p"
/protein_id="AAK89511.1"
/db_xref="GI:15159387"
/translation="MCGQNMPCPMPTVTEPMLHCSSERSATETIBGLPLADRLAKGIRLLISGELAPGOLSEAPSERFDPVSRNSLRALFTYDGLIRFANGVAVAVPMSVVIDIYRVRVVEGALRQAHKHPATKAMAAVEDAKITREEDMDVGVANMKFAHVVLDPSARISDPYERIAELRLCRGLKDPQLHSPFVDMNSKIISMBEGKAOBAENEBELYNLSERTVLKALEVAHEHTS"
2341..2955
/gene="AGR_L_1879"
2341..2955
/gene="AGR_L_1879"
/note="response regulator nast - Azotobacter vinelandii"
/codon_start=1
/transl_table=11
/product="AGR_L_1879p"
/protein_id="AAK89512.1"
/db_xref="GI:15159388"
/translation="MLRCVINADMTLRDLTIVIDENAIRASIEEGREAGYORVYIHEVGVATTELTLPDYIFIDLEPNNDMEHLEFQITRTYGRPIAMVYDRSDTASIEAAVEAGVAVVDGALKERVKPILDMAVSRFAERLQRELADASALEERVITERAKGILMKRGISEEERAPLQRTAMNEKKISIAQSVVTAAGILM"
2991..4286
/gene="AGR_L_1881g1"
2991..4286
/gene="AGR_L_1881g1"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_L_1881g1p"
/protein_id="AAK89513.1"
/db_xref="GI:15159389"
/translation="MAAPEKTSKSGEATQNRDADALVGSROKILRAGFIPLVDASVLIAAEFAPADREGLIDLVDXVSWANNRDLAPQDPIAHNLSPMPVASMGLGSPNPTTIPSLIRGNATITSTRLFAMKRLGLSEFAGALENARLALKLVLDREARGEAEPITLGMTVPSSHNVEFRYVLAAGGIHDDHVKLVVPPMTSDALADGFCVGAEMNIVYARNGVRIYAAKODLWPSAPERYIGKPRMASCOETVGRLLTALDAASMCDIADNDLSCGLADPRTYIGAPQSTIRVLAQERFIDQGNKRVIEKTFTRGHIANYPROSQSLMIVSQMRWGALSETGVNALAYRDIYRALGDKAKAGDDADIRIEQDGDRLFVDGFVPDPADIAGYVNSFAVRSVPSAPSGBA"
4544..5959
/gene="AGR_L_1886"
4544..5959
/gene="AGR_L_1886"
/note="nitrate-binding protein ntrA precursor, periplasmic - oscillatoriacean cyanobacterium"
/codon_start=1
/transl_table=11
/product="AGR_L_1886p"

gene
CDS
/protein_id="AAK89514.1"
/db_xref="GI:15159390"
/translation="MLPGTTVPVPAQTSPSGVFIFRRFRKRHRHPKWLKGNMRKK
IIFSGVSRRTLIKTATALVYATAPSSAFNAPFAPEYGAKLGFIALYTAAPLI
IAAGKGLPAKHMKEVEVYLKQASMGATADNLVLGASAGNDIGAHILTPMEYLMKTVGV
TONNYPVPTLITLANLDSOGISVAKVEYAEFGVQLDASKLKAAPKKAQKGLKAAH
TPPGTHDLMIRYMLAAGIDDPDQVSTIVVPPQMVAMVGVMDVPCVGPNNEDL
VNOGIGFTRACTTGELMKGHPEKALGMRADWTBKPNNATKALLAAMEAQOCDEMANK
BEMSTITLGRQFNPYPKDVGLRKLGNINYNGARVLENTGTQMLQMDWDSAPPHSHD
SMPIETIRMGCFAPDPTDVKLVAKVNNEDIRAAAKDLGVADLPASTSRKETFFPDG
KVPENSPSAYLESISIKAAAS"
6067..6566
/gene="AGR_L_1888"
6067..6566
/gene="AGR_L_1890"
/note="NITRATE TRANSPORT PERMEASE PROTEIN NRTB"
/codon_start=1
/transl_table=11
/product="AGR_L_1888p"
/protein_id="AAK89515.1"
/db_xref="GI:15159391"
/translation="MPMSALARKFNEBPQAPAVOKANVALSPHLPRIDLKRWGAAA
ANNTIPLVYLAIVLIGWQLCSAPGSSLPPSGVFESYDLIVSPFNHSGDIIIGAGL
LKLVLSEVAYAFGLAALAGIILGAVIGSGVAMRGDLGFQYLATVPYLAIPSLT
SAFQDSNBSATFVITPITSITWPIYINTAVGANINODRANVAKURLNLQPEFFKIMLP
SAPYITFGRLGIGLSMLAIVAAEMLTGVGIGIFPIDWANNSSRLPDIIIVALTIVG
VGFMLDKVLAALAGKVTITGAAN"
6980..7777
/gene="AGR_L_1890"
6980..7777
/gene="AGR_L_1890"
/note="NITRATE TRANSPORT PROTEIN NASD"
/codon_start=1
/transl_table=11
/product="AGR_L_1890p"
/protein_id="AAK89516.1"
/db_xref="GI:15159392"
/translation="MNAYLKLDHIDKPYDKGSGRAEVLKGINLTIEKEFVSVIIGSG
CGKSTMLNLIAGLITVSAGVAILENREYNEBPRAVWFQNHSLPMLSYVENNLAIV
AAVFNTKTRTERHDVNRNLDLVQMAHXDKKSEISGCKQKQVIGARALAMEPKIL
LIDPEFGALDALTTRAHLODAVMEIHARLGNMTWMTIHDVDEAVLSDRIWMTGSPAA
HIGELVDVPLARPRRIELASDKIYLKREAVLKFYERHRFVBAE"
7714..10230
/gene="AGR_L_1891"
7714..10230
/gene="AGR_L_1891"
/note="asimilitatory nitrite reductase large subunit PA1781
[imported] - Pseudomonas aeruginosa (strain PA01)"
/codon_start=1
/transl_table=11
/product="AGR_L_1891p"
/protein_id="AAK89517.1"
/db_xref="GI:15159393"
/translation="MPRCAGVPLRTSLPDRSRGVITMTQVLVINGMAAGRLLENL
FETAGCLDVTIFNAEPVAVDRITMLSVLSEKSEYDVIHNDERTAAENVTLHKKA
KYTRIDRANKVITSENGITASVDRVLVATGSLPTIIPVGHQLFGVLAIBDLDTVGM
LEIAGKGRALVIGAGLGLLEAYAGLKGQMDVIVHIMPTIMERQDPAAYVILEKA
INBRIDITTKANTKICLIGEEKVEGIELEDRRIIGKDWVAWVAPRASGILKAGGIA
VNRGVDDGQMDASIVYALGECGAEHMGVGVYAPLYESARLILIKENRIIGAVYGET
SYTRKTKVGTGINTLPSAGDPAGDDREIVLRADATYERKILIKENRIIGAVYGET
ADGSGPFEDLMKSTDISAMRETLIFGQAVQGSPLDEMAVAALPDAETCGCGVCK
GKITATVTSKGLTLDVRAHTKASASGNGCTGVQLMTLTLDITVNPAAVQPMCKC
TDLGHDVSRKLIKANGLKTIPAVMLKASGNGCTGACACRCPALNTYVLCDPEDEADCK
QSRVYENPHANIIODGTYSVVRPMGVTNVALEPAIADVADEFILPVYVGGAGRI
DLGILEKXDLPAVMADIGKAGIFSGQAAKGLATYKCVGQVOMCRPQDQSTGIRII
BEPFMGSGMTPAGLKAIVSGCPRNCEATCRQIGYICDSGCEIHFAPAAGDITGTV
LGLVTEDEALSHVIALTQMTREQARLYERIYKNAKRIGVDYERQIMDDADKKKAF
DRFVDSQKFAQVDFPMSERSVSGHDKRFPMAAIGFSEAA"
10227..10571
/gene="AGR_L_1892"
10227..10571
/gene="AGR_L_1892"
/note="asimilitatory nitrite reductase small subunit PA1780

```

{imported} - Pseudomonas aeruginosa (strain PAO1)
/codon_start=1
/transl_table=1
/product="AGR_L_1892p"
/protein_id="AAK89518.1"
/db_xref="GI:15159394"
/translation="MRTMSNDNIDIGDISDIPLGARCVYTPMGKIAVFAEENQVPA
IEBRCPKRGKGLSGIYHGASVYTCPLHNWYISLETGALGADGEVNTIVRNNDGRL
SIALETTIMTAAE"

```

	Query March	9.6%; Score 133; DB 1; Length 10676;	
	Best Local Similarity	50.2%; Pred. No. 1,9e-15;	
	Matches 572; Conservative	0; Mismatches 445; Indels 72; Gaps 7;	
OY	87 CGATGTCGGGGCAAGTTGSGAAGCTGTGAAGAAAGAAATTGTAATTCGGCTTTCATCGAAGCT	146	
Dd	4768 CGCTTTCCGGGCCACCCGACAGAACCGAAGTAGAAGGGTGCAAAAATCTCGGTTTTCATCGCGCT	4827	
OY	147 CACCGACATAGCGCGCGCTTGCGCGGTGGCGGCCTGAAAAAGGCTTCTTTCGAGGACGAGGGCCT	206	
Dd	4828 CACCGAAGCGCGCGCGCTTCATTCATTTGGCGCCGAGAAAGACTTTTCCCAAAGCATGCGCAT	4887	
OY	207 GTTGTGCAACTGTGAAGCG---CAGGCGCAACTGMAAGSTGTGATGATAGGGTCGTAA	263	
Dd	4888 GCCTGAGGTGAGGTTTCTGAAAACAGGCTCTCTGGGGGGCCACCGCGCACATCTCGTGCT	4947	
OY	264 TGGGAACTGAGCGGCTCGCAATGTGCGCGCGCGCGCTTGAAGCGGCGAGCGTTGCTT	323	
Dd	4948 CGGCGG-----CGCTCTGAACCGGCATTCGATGCGCGGCATATCTTGACGCGCATGCGCTTA	5001	
OY	324 CGGACCAAGCGCGCATATCGAGGGTCCGTTCAAGATGGGCTTCAACGGCAACCGGATTAAC	383	
Dd	5002 TCTGATGCAATCCGGCAAGGTGACGCAAAACAAGTCCGGTCCGATGACGATCTTCGC	5061	
OY	384 GGTTGCCAATAAATCTGGCATCGATGAGAAGCCGAACATACCCTGTCGAGGCGGTAAACC	443	
Dd	5062 ACGCTTATATCTGACAGCCAGGCGCAT-----TTCCGTGGCCAAAGAAATATGACAGAC	5115	
OY	444 GTGTCATCCGATCAAGGACATTTCTTAAAAACGGTGTGTGAAAAATCAAAAGCCGAGG	503	
Dd	5116 CGGGTCCAGTCCATGATCCCTCCAAGCTGAAGGCGGCTTTGAGAAAAAGAAAGGCCGACG	5175	
OY	504 CAAGCGGTTCAATATGCGCATGACCTTCCGGGCGGAGTCTCAACAATCAAACTGCGCTTA	563	
Dd	5176 CAAGGAATCAAGGCGCGCATGACCTTCCCAGGCGGACGACATGATCTTGTGATCCGCTTA	5235	
OY	564 CTGGCTGCGCGCTGCGGATTCATTCCTGGCTACTATTCGCGCGCGGACGACATTTCCGG	623	
Dd	5236 TTGCTGGCTGTGTGGCATTCGATCCGGAAGA-----	5269	
OY	624 CCAATATGCGGCGCAACGCTTGTGTGTGTGATACCCCGCGCGCGGAATATGCCGTTCAAGCT	683	
Dd	5270 -----GACGTCTTCACCAATGCTGTGTTCCGCGCGCGGAATATGTCGAAACAT	5316	
OY	684 GGAAGCGGGAACATTTTCGGTTATTTGCGTCGCGGACCGCTGGAACAGACAGGCGCTGTT	743	
Dd	5317 GAAGTGTGGAACATGAGACGTCTTCTGTGTATAGGAGACCGTGGAAACGACGACCTGTTAA	5376	
OY	744 CAAGGCAATAGGCGTGGCGGTGATCACCGATGAAGAACTTGGHAGGACACGCTGGAAAA	803	
Dd	5377 TCAGGGCATCGGCTTTTATCTGCTGTATCACCGGGGAACCTGGAAGAGGCAATCCGAAAA	5436	
OY	804 AGTTTGGGGGTGACCAAAACAATGGGGGGGAGAAATATCCCAACAACATATGCGGCGAGAC	863	
Dd	5437 GGGCTGTGGCATGCGCGCGCGCGCTGAGGAGAAAGAACCCCAAGCCACCAAGGCTTTTGCT	5496	
OY	864 CAAGCGCTGATTTCCGGCGCGCATCTGTGCTGGAACGCGCAATTAACAAGAACCGCAAGA	923	
Dd	5497 GATGGCGGTATGGAAGCGCAGCAATGTGTGACAGATGAGCGGAACAAGAGGAGATGTC	5566	
OY	924 AGCGATGAATATGTTGGCGCAAAAACAATACGTGTGATGACGTGAATGCTGTGCGGC	983	

[illegible]

gene
 CDS
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
 /codon_start=1
 /product="ABC transporter, substrate binding protein [nitrate]"
 /protein_id="AAL44713.1"
 /db_xref="GI:17742344"
 /translation="MKKIFSGTSSRRITLKTAAALVAVRTAPSPGAPAAVEPEVKGAKGPIALTDAPLIIAEKGLPAKMGPEVEVLKASWGTAPRNLVYGSNGIDGAHILTPMPLMHTGKVTQNNVPMTIILRLMDSOGISVAKIYATGYQLASIKLAERKADGKEIKAAATPPGTHDLIRIWLAAAGIDPDKDVSTIIVPPQVANKVGMADVCEGEPNBOILVNOIGTACTGELWKGHEKALGRADWVEKNPATLALMAVBAQOQDEMANKEEMSTILGRQMFVPPKDVGLKGNINYNGRVLENTGLOMKFQDDHASYEPHSDSWETENIRMGKFAPTDVKALVAKVNRDIWRAAKDGGVADLPASTSGKETFPDGKVPDPENPSAYLESLSIKAS"
 /complement(6423..7718)
 /gene="nrtC"
 /note="synonym: Atu3906"
 /complement(6423..7718)
 /gene="nrtC"
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
 /codon_start=1
 /product="ABC transporter, nucleotide binding/ATPase protein [nitrate]"
 /protein_id="AAL44714.1"
 /db_xref="GI:17742345"
 /translation="MAPEKTAGSKGBATQMDAPALVGSROKIIIRAGFIPLVDASVLIAAAFGADREGLILDVQVSMANVRRLAFQFDIAHMLSPMPVSMGLGNSPLSPITPSPGCGNAITLSTRLPARKMLGLSETAGALANALKILVLDMDARGEAPMLGMYTPSSHNEPERVYLAAGGHPDHDVLVVPMTSDALAAAGIDPGCVGAPMIVABEGVGRIVAKODIWPABEKVIGMPPEWASQGTGSLTLRLDAAASMCDLADNRDALSGALADPRYIGAPOSTIRRYLAGFSDSQGNRRVIEKFTFHGDHANYPRQSGSLWYISQMIKMGQALSETGVNAALSAVRPDIYRALDGDGKAPGDRIKQGDGDRFVDGFVPPADIVAGVNSPAVTSVPSAPSGEA"
 /complement(7754..8341)
 /gene="nast"
 /note="synonym: Atu3907"
 /complement(7754..8341)
 /gene="nast"
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
 /codon_start=1
 /product="two component response regulator"
 /protein_id="AAL44715.1"
 /db_xref="GI:17742346"
 /translation="MTLRDLTLVIDENAIRASIIIEGRLBEAGYORVTVIHEVNGARTIETLOPDVITDLENPNRDMELFOLRTGVRPIAMPVDRDPTASIEAAYGASAYVVDGLKGRKVPKILDMASRVAFSRRLRELDADKASALEERKVIIRAKGITLMKQGLSEBAFALLKQIMANEKKKISLTAQSVTTAAAGLM"
 /complement(8443..9138)
 /gene="Atu3908"
 /complement(8443..9138)
 /gene="Atu3908"
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
 /codon_start=1
 /product="transcriptional regulator, GntR family"
 /protein_id="AAL44716.1"
 /db_xref="GI:17742347"
 /translation="MLHCSERSAMTETTBGLPLADRLAKGIRTLISGLAQGISLAAFSERFDVSRNSLRFAFLTKDGLLRHEANRGVFAVPTMSSVIDIYRVRMVECEA

Query Match 9.6%; Score 133; DB 1; Length 10707;
 Best Local Similarity 50.2%; Pred. No. 1.9e-15;
 Matches 572; Conservative 0; Mismatches 495; Indels 72; Gaps 7;
 GALRQAMHKKPAIKMRAVEDAKIRREADWIGVSAANKFHAAVVDLDSARISDFYERIAAEHLRCFGLDPEQLHSPVDNMKSLIIMEEGAAQBAENEMELYNLSERT

Qy	87	CGATGTCGGGGCAGTTGGCAAGCTGGAAGAAAGATTGAAATTCGCTTCAACAGCT	146
Db	5941	CGCTTTCGGCGCCACCGCAAGACCGGAAGTGAAGGGTGCAAACTCGGTTTCAATCGCGCT	5882
Qy	147	CACCGACATGCGCGCGCTGGCGGTGGCCGCGGAAAAAGCTTCTTGAGAGACGAGGCGCT	206
Db	5881	CACCGACGCGCGCGCGCTTCAATTCGCGCGGGAAGGACCTTTTCGCCAAGCATGGCAT	5822
Qy	207	GTTTCGTCGCACTGGAAGCG---CAGGCGCACTGGAAGGTGTGATGATGATGGGTGCGTGA	263
Db	5821	GCTTGAGGTGAGAGTTTCTGAACAGGCTCTTGAGGGGCGACCGCGACATCTCGTGTCT	5762
Qy	264	TGCGCAACTGGAACGCGCTCGCACATGCTGGCGCGCGCGCTTACGCGCCAGCGTTGCGCTT	323
Db	5761	CGCGCG---CGCTCGAAGCGCATGATGGCGGCGCATATCTGACCGCGCATGCCCTTA	5708
Qy	324	CGGCAACGAGCGCGATATCGAGGTGCCGTTCAAGATGGGCTTCAACGCGCAACGCGATTAC	383
Db	5707	TCTCATGCAATACCGGCAAGGTGACGCAACAACTGCGCGGTCCGATGCGATCTCCGCG	5648
Qy	384	GGTGTCCAAATGAATCTGGCATGATGAAAGCGGAACATCCGCTGGAAGCGGTTAAAC	443
Db	5647	ACGCTTAACTCTGACAGCGCAAGCGGCAAT-----TTCCGTGCGCAAGGAATATGCAAGAC	5594
Qy	444	GGTGCATCCGATCAAGGACGATATATATTAACCGGTGCTGCAAAATATCAACCGCAAGG	503
Db	5593	CGCGGTCCAGCTGATGAGCTCCCAAGGTGAAGGGGCTTGGAGAAAAAGAGCGGACCGG	5534
Qy	504	CAAGCGTTGAATATATGCGCATGACTTCCGCGCGGATCTCAACAATCAACTGCGCTTA	563
Db	5533	CAAGGAATCAAGCGCGCGCATGACTTCCCGCGCGGACCAAGATCTCTGATCCGCTTA	5474
Qy	564	CTGGCTGGCGGCGTGGGGGTATCAATCTGGCTCTATTTGGCGCGCGGACATTTCCGG	623
Db	5473	TTGGCTGGCTGGTGGGCGATTCATCGGACAAAG-----	5440
Qy	624	CCAATTCGGCGGACAGCGCTTGTTCGTGTAACCCGCGCGCGCAATATGCGTCAAGCT	683
Db	5439	-----GAGTCTCCACCAATTCGTCTTCCGCGCGCGCAATGCTCCGAACAT	5393
Qy	684	GGAAGCCGCGACCAATTTGCTTATTTGCTGCGCGAGCGGTGAACCAAGCGCGCTGTT	743
Db	5392	GAAAGTCGGCAATGAGAGCTTCTGTGTGAGGCGAGCGGTGGAACGAGCAGCTCTTTAA	5333
Qy	744	CAAGGCAATAGGGGTGCGCGGTATCAACCATGAAAGACTCTGGAAGACACCGCGAATA	803
Db	5332	TCAGGCGATCGGCTTACTGCTGTACCAACCGGGAACCTTGGAAAAAGGCGATCCGAAAA	5273
Qy	804	AGTCTTCGCGGTGACCAAAACAAATGCGCGGAGAAATACCCCAACCTATCTGCGGTTAC	863
Db	5272	GCGGCTTCGCGATCGCGCGCATGCGGTGAAAAAGAACCCCAAGCCACCAAGGCTTTGCT	5213
Qy	864	CAAGGCGCTGATTCGGGCGCGCATCTGCTGACCGCGCAATTAACAGAACCGCAAGGA	923
Db	5212	GATGCGCGTGAATGAGAGCGGACGATGTCGACGAGATGTCGACAAAGAGAGATGTC	5153
Qy	924	AGCCATCGAAATCTGCGCGCAAAACAAATACGTGCTGCTGAGATGCTGCGCGCG	983
Db	5152	CACCATC-----CTGCGCAAAAGCGATGTTCAACGTGCCCGCAAGAGATGTTCTCGG	5099
Qy	984	TACCATGAACGCGACTTTCGATACGAAAAAGACGATTAACGCGCGCTACCGGACTT---	1040
Db	5098	CGGCTTCAAGGCGCAATATTAATTAACGCAAGCGCGCGGTGCTGAAAAATATCTGGCTGCA	5039
Qy	1041	CAACACCTTCTTTCGCGACGCGCGCATATCCGCTCTTACAGCATGCGAGTCTGATCT	1100

Db 5038 GATGAAGTCTGGCAGGACCATGCTCTCTATCGGTTCCACAGCCATGACAGCTGTTCAAT 4979
Qy 1101 GACCCGCTGAGGCGCTGGGGCATGATCAATGATTAACCGGACACTGGTATCTGGA 1160
Db 4978 CACCGGAAACATCGCTGGGGCA-----AATTCGCGCCCGGATACGACCTAAAGGC 4928
Qy 1161 TACCGGACGAAGCTGTACCGCCGACATCTATCTCGCGCGGCGCAAGGAACTGATCG 1219
Db 4927 GCTGCTTCCGAGAGCTGAACCGCAGATATCTGCGCGGACGCGCGAGAGATCTCGCG 4869
RESULT 9
LOCUS Ap003000 348077 bp DNA linear BCT 15-MAY-2001
DEFINITION Mesorhizobium loti DNA, complete genome, section 7/21.
ACCESSION AP003000 BA000012
VERSION AP003000.2 GI:14023048
KEYWORDS
SOURCE Mesorhizobium loti
ORGANISM Bacterii; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Phyllobacteriaceae; Mesorhizobium.
REFERENCE
AUTHORS 1 (sites)
Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S.,
Watanabe, A., Ideasa, K., Ishikawa, A., Kawashima, K., Kimura, T.,
Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A.,
Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M.,
Takeuchi, C., Yamada, M. and Tabata, S.
TITLE Complete genome structure of the nitrogen-fixing symbiotic
bacterium Mesorhizobium loti
JOURNAL DNA Res. 7 (6), 331-338 (2000)
MEDLINE 21082930
PUBMED 11214968
REFERENCE 2 (bases 1 to 348077)
AUTHORS Kaneko, T.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research; Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail: kaneko@kazusa.or.jp,
URL: http://www.kazusa.or.jp/rhizobase/
Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)
On May 11, 2001 this sequence version replaced gi:11994975.
COMMENT
FEATURES
source
1. 348077
/organism="Mesorhizobium loti"
/mol_type="genomic DNA"
/strain="MAFF303099"
/db_xref="taxon:381"
complement(169..1065)
/gene="ml12551"
complement(169..1065)
/gene="ml12551"
/codon_start=1
/transl_table=11
/product="probable transcriptional regulator"
/protein_id="BAB49656.1"
/db_xref="GI:14023049"
/translation="MALDWDKLVFPAALAAAGSPTAAATLHLSQASISROYSALEHD
VAVPLFNHARGLVTEGEMLPFAHDVIAKLKLTETKDRPSGLAVTTVG
LGAGWTERVOSEFIEIPEISQILANEELKLTWRQADALRLPOOPDLIORLF
TVPHLYAASVYAKHGRKPAVSSELRNRIYFGLPYPSHLSKMLTETVDFEGGQR
VSLQINDLSTIKRAVOGAGIAMLDPVYINKDSGLVQLPTEVPSPTVYAYPDAM
KNQAKLHVRFDITIAKARWSF"
complement(1156..2130)
/gene="ml12552"
complement(1156..2130)
/gene="ml12552"
/codon_start=1
/transl_table=11
/product="thioredoxin reductase"
/protein_id="BAB49657.1"
/db_xref="GI:14023050"
/translation="MTTTHA PVI.IIGSGNAGTYAVARAMLKPLVAGLQCGQLM
ITTDVBNYRGPDPYIQGPIVMEQMKQAEHVTDDIINDITTEVDLNVPRFKAGDSGT
TTTADALITATGAQAKWIGIPTEDPFMGVACATCDGFYRGGDVAVAGGNSAVE
EALYLSNLAHSVTVIHRSDFAERILERRLLQKNVRYIWDVDEITGRGKELP
PVEGIGKLGAVTGAETHLKVDGVVVALGHAVALFVYKLRKXKPGYLMTAPNSTRT
DVGVAAGDVDDVYRQAVTAAAGCMAALAEKVIAGIEVHREAAE"
2378..2857
/gene="ml12554"
2378..2857
/gene="ml12554"
/codon_start=1
/transl_table=11
/product="probable transcriptional regulator"
/protein_id="BAB49658.1"
/db_xref="GI:14023051"
/translation="MPLKNADLDAIDMKILRELQADGRMTVELSNRVGISAPPCRLRV
KLIERAGIIRGYRALNAPALGDVVAFLVGLHQAQDELTFPARETGMPVIRAW
MVSGRSDFLHCTASDLGAFQFTVEILSTFNVDTVRLTIRRVKDEGLVSFRS"
3093..3887
/gene="ml12557"
3093..3887
/gene="ml12557"
/codon_start=1
/transl_table=11
/product="glucosyl transferase"
/protein_id="BAB49659.1"
/db_xref="GI:14023052"
/translation="MTRFASPVSAIICVNEADMTGPCLESIDPCAIIIVDSGSDTG
TIECVSYIAKGYPIRLHNDWPGPPROFSLDIAIOTPCWCSIDPEDEVDQLQSI
VAITQADDKISGYTIRRDYMKGYVAHWLVNMLPFRREGATIDLTARVHAF
OYRGRTGDEGVILHRRRILSYEDGLAANAVSSIKAVTLVYKRGKPGIVRLVLSPLG
NPLKYLAKRYFLCGRHGFVMSVMYIISFATEALYEASNDPV"
complement(3903..4556)
/gene="ml12559"
complement(3903..4556)
/gene="ml12559"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB49660.1"
/db_xref="GI:14023053"
/translation="MADKTGPHPLTPRGWVIERIGGRVFPPLPALFLDRDGTINVD
TDPSDPAEIVRPMQLPALATANAGIPIVYVNTNOSIGARVFGMSAPAAVGVLE
ILREBGRVNDWTLACATHEAGVGLATIDHPRKRNPMGLZEAGKRLDLDRSLIVG
DKLADWQAKRAGLAQGMVDEGAIVQGFALRPLRDSSELDLALAEITGRNRS"
complement(4549..5121)
/gene="ml12561"
complement(4549..5121)
/gene="ml12561"
/codon_start=1
/transl_table=11
/product="probable phosphohexose isomerase"
/protein_id="BAB49661.1"
/db_xref="GI:14023054"
/translation="MSELNDYVRSAAITMVERDLTGMERASAVVTLSSGKL
LIAAGGSASDAIHITAGELVGRFLERKAYIVIALPANAATITANGNDYGPPTVSRO
VEAHGTAGVGLAISTGNSPSITAAQARMDMTVIGLTGDTGKPLCDILLLNV
PSTSTPIIQOGLCDYHHLCEVEARLNG"
complement(5129..6619)
/gene="ml12562"
complement(5129..6619)
/gene="ml12562"
/codon_start=1
/transl_table=11
/product="ADP-heptose synthase"
/protein_id="BAB49662.1"
/db_xref="GI:14023055"
/translation="MTKNPPSPPEPLHRAIRFGGVTVLVGDLILDRFVNGVIRIS
EAPLPIVARGGETSAMGAGNVANIVLSKARALPVSDITGDTAGLIRHRAALAGAI
ETAGLSQORGRMTSEKSRFSAIINOVLVDFDEIEIPLDETETAGLIRHRAALAGAI
VILSDYGGIILDDVVAELIACREAGRPVVDIGRGRYVAGATATIPNKEIGEA
VGHAVFADDEIVAAARELISAHGPFVAVTSSEKMSVYRGDEAHHTAINTQAREVDS
GAGDVIATFPALALASGADPVAAASINAAAGVYVKGKGTARTLTVEILTGLFRSHGP

FEATURES	source
PUBMED REFERENCE AUTHORS	12597275 2 Kaneko, T., Nakamura, Y., Sato, S., Minamiaeawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Ideawa, K., Iriyuchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpō, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
TITLE	Complete genomic sequence of nitrogen-fixing symbiotic bacterium
JOURNAL	Bradyrhizobium japonicum USDA110 (supplement)
MEDLINE	DNA Res. 9 (6), 225-256 (2002)
PUBMED	22485002
REFERENCE	12597279
AUTHORS	3 (bases 1 to 300150) Kaneko, T.
TITLE	Direct Submission
JOURNAL	Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: kaneko@kazusa.or.jp / rhizobase / .URL: http://www.kazusa.or.jp/rhizobase / Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
FEATURES	Location/Qualifiers 1. 300150
source	/organism="Bradyrhizobium japonicum USDA 110" /mol_type="genomic DNA" /strain="USDA110" /db_xref="taxon:224911" complement (14. .862) /gene="b112729" complement (14. .862) /gene="b112729" /note="ORF ID: b112729 probable ABC transporter permease protein" /codon_start=1 /transl_table=1 /protein_id="BAC47994.1" /db_xref="GI:27350986" /translation="MDRNRERKMTPRQIGKIGLMSLVIVSPAILFELMMLSLSLK FEIDNAPVPVPEPEHIAWKNYADVLSNRELYFINSLVVTGSATLALVGPAGY GIDAMAAKSAIVILLIARIITPGLSYLIPLFLFQMLGLGLIVQIILHVVTPIV WIMYGPEETPELEBAALIDGATRCQVFRVHALPIAKPGIAVFTILAVISMNPFV GIYAGERTRLPVAVYNMISFPDLSNCPPLAALIVTLPVLLTVFAQRQIVAGLTR GAYKGC"
gene	complement (835. .1773) /gene="b112730" complement (835. .1773) /gene="b112730" /note="ORF ID: b112730 probable ABC transporter permease protein" /codon_start=1 /transl_table=1 /protein_id="BAC47995.1" /db_xref="GI:27350987" /translation="MSALITQADPAARSDAAREKLAPSPSPWPVVALVYVLAIIIF PWFTTMSLMEKVGSPITFVGLSNYLRLLSDPRLEAGHTWVYALSVLPVLVIG TFPAVVRHOKPAGRGFLRGVFIEMPMATPVALAVMTMFPDGLGVNTYILSLGIPA QLVAVFPAVTPSLVLTVEQWTEPLVNLVILGIGLAALPTERYESAQIDGANLMQFFER ITLPLIMPLFIAGIMIMIDAVKSFCDIIFAITGGPGSASPTINIYLYSAVFTYDGG YGSIAVVFPLIYTLAALVWLVAQRQVWTEIASDA"
CDS	complement (1770. .3152) /gene="b112731" complement (1770. .3152) /gene="b112731" /note="ORF ID: b112731 probable ABC transporter substrate-binding protein" /codon_start=1 /transl_table=1 /protein_id="BAC47996.1" /db_xref="GI:27350988"
gene	/translation="MARSLVSLPKTMSHREDPEVDYVIRSLTLLGTAAAGLSLTG LPAAPVNMKKYACTKLIVLAKPGRDGNQKIKETELTGLIYEAERQIPEQQROR VYIEFTSGRPSFDVWLSIVQKQFBEKAGLADMTPEKDPSTLADLVESDFSAGG LQAKAGDQGLSLPMSVDVFILYYNRELQKGVAVPKTLDENVAAEKTLPKEGTG YGVAGRGTLRANNAALMTNFIYNGSEFLDCKRNLITDGPALATKLYQTLTSGAPP
CDS	

GYAGENMMESMASFTQGRSAMVIDGVGNAPPLEDPASRVVGVKGYTVVPAGPKGHYS
 STYVDGIQGLAASKNKEAAYILLQWVVSQKQARLLQGGVPPRSSILNDPEIQKG
 KNPPEQLQGVIESGKISKLGILPVVVPVVAEPRFIVGALLTATLSGADPAELKAKNDQ
 RPLRSEKA"
 complement (3179. . 4012)
 /gene="b112732"
 /complement (3179. . 4012)
 /gene="b112732"
 /note="ORF ID:b112732
 putative cytochrome c"
 /codon start=1
 /transl_table=11
 /protein_id="BAC47997.1"
 /db_xref="GI:27350989"
 /translation="MDRRQRLFVHRLGIDHGRMKFWRAIYAVTLLAAPALLAGEIAR
 EARRSGYAPMGSDTPADODDIPANQGMFLVLDGEALFTKKGSAEKACADQSGARRG
 MGVARVYAPAEKAVGRPTLLDQRIINLGRAHQQATPELPSRRLLIASAVAHOSRG
 VAITGADDDPOLKQFVQGRBDLFWQREGLNLTACTGHNDNDRHLGAPITTOAQGTG
 PLYRLBWTQSLERRLRSCMSGVAQYVDYGAPELVLELYMSRAGLMEETPAVAV
 P"
 complement (3955. . 4272)
 /gene="b112733"
 complement (3955. . 4272)
 /gene="b112733"
 /note="ORF ID:b112733
 probable sulfur oxidation protein"
 /codon start=1
 /transl_table=11
 /protein_id="BAC47998.1"
 /db_xref="GI:27350990"
 /translation="MAALINVPAKARRQDVIEIRTLTSHMETGFRHTADGRLVPRDII
 ITSFICRYNGTEVFRADLFPAILANPYLSPFTVAKESGKPEFEWTDGNGYSTASASII
 TYE"
 complement (4272. . 4739)
 /gene="b112734"
 complement (4272. . 4739)
 /gene="b112734"
 /note="ORF ID:b112734
 hypothetical protein"
 /codon start=1
 /transl_table=11
 /protein_id="BAC47999.1"
 /db_xref="GI:27350991"
 /translation="MPTTRQFLSLAGGTTAAGTIPVTLRLPLEATPAMLTAIRNVV
 GSAPRTIRQVADTIPLVENANTVPMTISVASPMTPDDYKSIHVFNKSNQPIGNF
 HIGPAGNAGVQSTRIRLADTQKVAIALSDSFQVADVVYTLAACTEEMN"
 complement (5385. . 6665)
 /gene="b112735"
 complement (5385. . 6665)
 /gene="b112735"
 /note="ORF ID:b112735"
 /codon start=1
 /transl_table=11
 /product="flavocytochrome C flavoprotein subunit"
 /protein_id="BAC48000.1"
 /db_xref="GI:27350992"
 /translation="MRPNAPVYTRRNVLGITTAATSAVPSILRAQAGRVVVVGGG
 FGACAAAILRANKLQVTLIPNKVAILSPFSNEVIAGIRETEAOFGVDKIVAG
 ITVAQAATYTPDEPRSTYTTDGGATITYTDRLVSPGIDDFRYEALPGTDEAASEMPPH
 MWAQQTILLRQLSEADGGTVAIAIPANPSRCPAPYERASLIHAHLTKTKRRADV
 LILDKDSFSOORLEFARMKELYGNMTERIGLSQGVSTVSDASTKTIYTFQGYTDD
 VANVLPQRAQHLAQIAGAADATMCPIDPTVFESKLVPMNIHVIGDACLGGGIPKSMAS
 AASAGKCAKAIIVNLAGRAPETPRLTGVCTTVAPGYGFSLAGNVQPRGDIYAEVSE
 GGANSVPVAPPEIARREADEARNFQITTDFTG"
 complement (6782. . 8968)
 /gene="b112736"
 complement (6782. . 8968)
 /gene="b112736"
 /note="ORF ID:b112736
 putative aldehyde dehydrogenase protein"
 /codon start=1
 /transl_table=11
 /protein_id="BAC48001.1"

RESULT 13
LOCUS AP005951/c
DEFINITION Bradyrhizobium japonicum USDA 110 DNA, complete genome, section 17/31.
ACCESSION AP005951 BA000040
VERSION AP005951.1 GI:27352606
KEYWORDS Bradyrhizobium japonicum USDA 110
SOURCE Bradyrhizobium japonicum USDA 110
ORGANISM Bradyrhizobium japonicum USDA 110
REFERENCE 1
AUTHORS Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Ideesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimo, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
TITLE Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110
JOURNAL DNA Res. 9 (6), 189-197 (2002)
MEDLINE 22484998
PubMed 12597275
REFERENCE 2
AUTHORS Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Ideesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimo, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
TITLE Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110 (supplement)
JOURNAL DNA Res. 9 (6), 225-256 (2002)
MEDLINE 22485002
PubMed 12597279
REFERENCE 3
AUTHORS Kaneko, T.
TITLE Direct Submission
SUBMITTED (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; 2-6-7 Kazusa-kametari, Kisarazu, Chiba 292-0812, Japan
URL: http://www.kazusa.or.jp/rhizobase/,
Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)
LOCATION/Qualifiers
1. 298700
/organism="Bradyrhizobium japonicum USDA 110"
/mol_type="genomic DNA"
/strain="USDA110"
/db_xref="taxon:224911"
48. 512
/gene="b1r4344"
48. 512
/gene="b1r4344"
/note="ORF ID: b1r4344
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC49609.1"
/db_xref="GI:27352607"
/translation="MTQATPASTTPKRRPRPQVNLTDAAQRISSELTQRADSEIVG
LRGVNKGCGAGQSYVEYAHVRPTDEVEDKGVILVDPKAVLFLIGHRDLGRO
DAGPVRLQOPGGLRLRLRRIGRAAGAEEDRVRGVALPFRSAVVLAKARTHTA"
584. 1069
/gene="b1r4345"
584. 1069
/gene="b1r4345"
/note="ORF ID: b1r4345
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC49610.1"
/db_xref="GI:27352608"
/translation="MGPGFODDEBERGAMDRRELLIDLPADFGPYTIRKMFSGYISA
DGTNPLALRAGLFLRADELTTIDPETEGSKRFQYSTRAKTYTNSVTLPARLPDDSD
GELAQMARALAAAGRAKYKQRPKQKASKKYAPKKPVAKRRPAKKAARVYKAGAH

gene
CDS
R
complement(1134..1664)
/gene="fabA"
complement(1134..1664)
/gene="fabA"
/EC number="4.2.1.60"
/note="ORF ID: b1l4346"
/codon_start=1
/transl_table=11
/product="3-hydroxydecanoyl-[acyl-carrier-protein]
dehydratase"
/protein_id="BAC49611.1"
/db_xref="GI:27352609"
/translation="MPNRPDFTPTPQSYTRELLASSSEGGFPGCAOLPAPNMLMD
RISBISLDGEGKCHIVSELDIVGHWFFDCHFGDPLMPSLSGLDAWQMTGVLG
WSGSPKQKALIVGEVCELTPTVQVRVYVAMRMVRGLVIGIADGRVADGAC
VFTAKDMRGVLTAKAVE"
complement(1763..2830)
/gene="b1l4347"
complement(1763..2830)
/gene="b1l4347"
/note="ORF ID: b1l4347
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC49612.1"
/db_xref="GI:27352610"
/translation="WVKVLDHERETMAFAEVALGQIRSLRQTAIPRNVEIWWYATGY
NAPLNKININETLARKKLTLEADLEQIYETVLSHITTDIDKAGRVGEIDVWKL
GVALMTGSDYDASLGATRLSSAKSREQIKAVETLVSTSEMETNVALDRLLTS
RNEISNLOOSLEAIRABSLTDLPLGINKRKYDPMIGMAVOSLAGSEPLISLLPID
HFKSPNDSYGHITGQVLRVLSKQRTIKODITARRGSEFVILLPTAARQALTV
ADHRAVWAKELKKKSTELIGRVTVISVSMKRPDDTDLTERADCLYAAKRG
NRVICEADPEFAVETHSRVA"
complement(2994..4496)
/gene="rhlE"
complement(2994..4496)
/gene="rhlE"
/note="ORF ID: b1l4348"
/codon_start=1
/transl_table=11
/product="ATP-dependent RNA helicase"
/protein_id="BAC49613.1"
/db_xref="GI:27352611"
/translation="MSFSNLGSEKVLAAVAATGTTTPPIEOQAIPIYLAKQVIGI
AQTGKXTAAFLVPLMTLLEKRRARMRPTLLEPTRELAQVKEENFVRGAGQNL
VALLIGVSPGQDAKLTGVDVLIATPGLDHTERGLLTVGLVLIIDADMLD
MGFIPDIERICKLVEPTROTLEPTAMPPEIRLITETFLNPOKYEVSQKATVAVT
OSQVPAKKAHEKRELLRLREAKDLKNAIIFCNRKRVAVIHKSLQKSPVGLH
GMDQPARMALEQFRKSELPLVASDVAAAGLDIPVSHVNPFPVPHRPDYVRVG
RTGARGSGTALSTVPLDQKSWAIEKLIGSITRAEGDYVHAADGATDPRQR
GNERSRGGRGKPRGRDRSHRPEARPSARAPAPAPRPRPRPREARPSPEPR
HGSRRQANASHPVSGRPEPRRADTEPCDHSLLPFLRPVSPAGA"
complement(4787..5782)
/gene="b1l4349"
complement(4787..5782)
/gene="b1l4349"
/note="arac family
ORF ID: b1l4349"
/codon_start=1
/transl_table=11
/product="transcriptional regulatory protein"
/protein_id="BAC49614.1"
/db_xref="GI:27352612"
/translation="MYRWCTDEVEPOERFDYRREVRAGLFGVTABLERERRADPFE
PSLRQAGGLVELKASPYTVRSSTDIATPGDSICVYQOLSGSGWGMGMSPAI
ANGSRATSHTDLPRTATLADGAFRLIKLPVSRIPRQDRIRRLTETPTDHALP
LIAACFADGEVNAAGDGDGADATSLVQTSIVQALAHALTERGIVRGSRIGQ
ALARTARLSQARLLIARQIONDLAPAVADLIGSVRLHMLFPAEKSFSQTYVDE
RLKRSRLRLREPERLLIADIATSCFESLATTYRVNAAYGMAPDIFRVRVEGH"
5935. 7503
/gene="b1r4350"
5935. 7503

gene
CDS

```
/gene="D1r4350"  
/note="ORF ID:D1r4350  
hypothetical protein"  
/codon_start=1  
/transl_table=1  
/protein_id="BAC49615.1"  
/db_xref="GI:27352613"  
/translation="MDCTRGAGKRSKCSNGNSMTRLRLRIILVYTLGLAASLGGRFLPAC  
AKERALVGNNDYRNVPFLKLAVNDARTMDTTLKQIGSPVMAVSNSSROODFSSTLLA  
FDRLAEPDQTAFFPYAGHGFELIAGONTYLLPTDVPAAETGDEELVRDAS ILADRIVTEL  
ONKKRSTAILLVFDCARNRPFERKGTBAVAGGGGALA PMMQLEGVFSVSGSPRTADL  
RLSNDADPNVSVFTRTPAKKELLOPGEINLVQVQRRLRVSEMAITVHQRQVYFPDQM  
VDDVLGSAAGDAARAPADPPKQALALPVSVVRVPEKFTTNA PLSFSGHNGSVV  
VSPEDPLTGLISMRAGAGDFRRTGFTITDLPTRTKRMPNPSITLPPDAQLTIEVRY  
VDQSGMDGPPPIRFDPPEALILRQDKLLDMATATSWLSFRFENGLVLYYTHLVSYRA  
IREVIGIDTAVPQNLVMPGCDMDRDPISAITAGMPLTWKLA PATQFVSVELVLYDGSV  
SEIKSFRAANSNN"  
7987. 9162  
/gene="D1r4351"  
7987. 9162
```

gene	9644, .10108 /gene="b1r4352"
CDS	9644, .10108

/note="ORF_ID:dlr4352
hypothetical protein"

gene
complement (10205, .10975)
/translation="MFGEQGITISLDKNGSLNTAFAGADVDVSRANVLTATVVGKAG
FAVQNNLPLYIKGAYAGVNNRLSVSDVTGVPAPGSGQTHMANGMTVGAGMEYGVTRNVI
VGLAEVAFSGQTFVLGGSTTNTYTFDAKPRDLQMAVVRASVKEDDAVILARY"
complement (10205, .10975)

```

/codon_start=1
/crystal_table=1
/product="2-deoxy-D-glucuronate 3-dehydrogenase"
/protein_id="BK049618.1"

```

Query Match	8.1%;	Score 113.2;	DB 1;	Length 298700;
Best Local Similarity	54.9%;	Pred. No. 1.2e-11;		
Matches 267;	Conservative 0;	Mismatches 213;	Indels 6;	Gaps 2;

131 TGGCTTCATCAAGCTCACCGACATGGCCGCTGGCGGTGGCCGCGAATAAGGCTTCT 190

DB 263952 TCGGCTCATTTCCGCTGGTTCGATGCCGCCGCCGCTGATCGTCCGCCGTTCGACACAGGCTTTA 2638

[illegible]

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

Db 263832 ACAAGCTCAATACGCCCTGTCGACGCGCGCCCAATCTGCTGCGCGCCCGTGGATCGCGT 2637

[illegible]

RESULT	14
AE007223/c	
LOCUS	AB007223
DEFINITION	Sinorhizobium meliloti 12129 bp DNA linear BCT 03-JUN-2004
ACCESSION	Complete plasmid sequence.
VERSION	AE007223 AE006469
KEYWORDS	AE007223.1 GI:14523386
SOURCE	Sinorhizobium meliloti 1021

CITATIONS

Barnett, M.J., Fisher, R.F., Jones, T., Komp, C., Abola, A.P., & Barnett, M.J. (1987) Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium. 1 (bases 1 to 12129)

REFERENCE AUTHORS

TITLE Gunjal, M., Hong, A., Huizar, L., Hyman, R.W., Kahn, D., Kahn, M.L., Kallman, S., Keating, D.H., Palm, C., Peck, M.C., Strzacki, R., Wells, D.H., Yeh, K.-C., Davis, R.W., Federspiel, N.A. and Long, S.R. Nucleotide sequence and predicted functions of the entire *Saccharibolus meliloti* pSymA megaplasmid
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9883-9888 (2001)

PUBMED 11481432
REFERENCES
AUTHORS
2 (bases 1 to 12129)
Barnett, M.J., Fisher, R.F., Jones, T., Komp, C., Abola, A.P.,
Galay-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J.,
Bujal, M., Hong, A., Huizar, L., Hyman, R.W., Kahn, D., Kahn, M.L.,
Katman, S., Keating, D.H., Palm, C., Peck, M.C., Surzyski, R.,
Wells, D.H., Yeh, K.-C., Davis, R.W., Federspiel, N.A. and Long, S.R.
Direct Submission
Submitted (23-MAR-2001) Biological Sciences, Stanford University,
371 Serra Mall, Stanford, CA 94305, USA

```

FEATURES
    source
        Location/Qualifiers
            1..1212
                /organism="Sinorhizobium meliloti 1021"
                /mol_type="genomic DNA"
                /strain="1021"
                /db_xref="taxon:566834"
                /plasmid="pSymba"
            complement(103..318)
            locus_tag="Sma0580"
            complement(103..318)
    CDS

```

```

/locus_tag="Smau580"
/function="Miscellaneous; Unknown"
/note="glimmer prediction"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

```

gene
CDS
/protein_id="AAK64962.1"
/db_xref="GI:14523387"
/translation="MSSRALQKSTVTRGGLGQFWAVAKASFRANSDSSRTACPDGAIKG
ALVGAATKPPRSRQTSALQIADGFPVNG"
/complement(394..1194)
/locus_tag="SMA0581"
/complement(394..1194)
/locus_tag="SMA0581"
/function="Cell processes; Transport of small molecules"
/note="glimmer prediction Excellent match (e-87) to
nitrate transport protein ATP binding component nasd from
Klebsiella oxytoca starting at amino acid 35 of glimmer
prediction. Met33 of glimmer prediction has good SD
sequence and was called SMA0581 start. Chrd, Nasd and NtrC
are synonyms. Possible operon with SMA0581, SMA0583 and
SMA0585"
/codon_start=1
/transl_table=11
/product="Nitrate transport ATP binding protein, probable"
/protein_id="AAK64963.1"
/db_xref="GI:14523388"
/translation="MTKSYLSLELDKSFERGQRTVVKQVSLTVDKGEFISITIGHS
GGKSTLNI VGLTQATGVVLDGKVVDEBGPDAVVFQNHSLPMLTVYENRILA
VYKVSRTNKGREHWTRNLELVQMAAAEKHPSEVSGKQGVGIRALAMEPKV
LLDEPFGALDLTRAHLQDOVMOIATGNTVMTHTHDVDAVILGSLDRIVMTMGPS
ARVGRIDVPLARPRRIELASDRITVLTRESVAKTLYERHFRVNAE"
/complement(1208..2101)
/locus_tag="SMA0583"
/complement(1208..2101)
/function="Cell processes; Transport of small molecules"
/note="glimmer prediction Possible operon with SMA0581,
SMA0583 and SMA0585. Excellent match to nitrate transport
proteins in Klebsiella (NasB) and cyanobacteria (NtrB)"
/codon_start=1
/transl_table=11
/product="NtrB, Nitrate transport permease protein"
/protein_id="AAK64964.1"
/db_xref="GI:14523389"
/translation="MSVNLKUKPQATPOTPAQVIALGOTASRGIDGRLSRPTQTVT
NLPLVLITPFLTAMQLICSSPESLIPPSRVLESWELIAPFIQGVQVQGFEMH
VVASLQRVAVGAMAAGVAVLGLVGSALAMRGDPIFOVLRTPPLAMVLSIAA
PDGGPSAIFVFIETAIWPIIITNTAVGIRINIPODYONVAKVRLINSFEFRGIMLPAA
APYFTGARIGIGLSWLAIVAAEMLIGVGIGFPITDAMNSSLISDIYALIVGIVE
FLDLRLALVGAATVGTANA"
/complement(2164..3456)
/locus_tag="SMA0585"
/complement(2164..3456)
/locus_tag="SMA0585"
/function="Cell processes; Transport of small molecules"
/note="glimmer prediction Excellent match (e-143) with
nitrate binding protein precursor proteins from several
cyanobacteria. Possible operon with SMA0581, SMA0583 and
SMA0585"
/codon_start=1
/transl_table=11
/product="NtrA-type periplasmic nitrate transport binding
protein, probable"
/protein_id="AAK64965.1"
/db_xref="GI:14523390"
/translation="MTKTLFGGLTRRSVLTQTATAWVGAAARTLLPSGAPAGAPET
AKATIGFALTDSAPLIITAKEKGLPDKRGMTVEVYVKAQSMGTTDNLVLGSAAGCID
GAHIIIPMPYLLISTGKVTONQPLPAILIARLNDQALISGAAVADLKVIGIDASVLK
DAFAKKCGGEAAKVAATPPGGTHDLWIRYMAAGIDPDQVETIIVPPQPMVAMK
VGTMDQCFVGEPMNQLVQKIGYAVNTAAIIMASHPEKSPFARMADWAKNRAVAL
VVAIEBAQWCDMDANKDELAIKIVGKSWFNVPPIQVDRILKGEYDYGKQIVENSPH
FMKPRERIASVPSHDAWELTENIRMGKLAADPTDKGLIAVYNNEDIWREAKDLGI
SDIPASTSRGPEPTFPDGKVFDPANPEYTLKSLAISRIA"
/complement(4274..4732)
/locus_tag="SMA0590"
/complement(4274..4732)
/locus_tag="SMA0590"
/function="Miscellaneous; Unknown"

repeat_region
gene
CDS
/note="glimmer prediction"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK64966.1"
/db_xref="GI:14523391"
/translation="MTCGPYHDSVVRPAAACAGCLSYLSDPTGRGRLPGEAQAPVA
RHVSPARBGCTDDGSGGTAATPNNPRITLASQTOAQDSQHRQERTTSLSRHSRQ
SGRQRAAFALITAPHPSMARRGRAPVAMAAVATXTPPRGVCALWPS"
/complement(4518..4957)
/standard_name="ISRM17"
/note="SMA002; predicted by homology, possible ISRM17
fragment, 92% identical over length to ISRM17"
/evidence="not experimental"
/tpc_family="ISRM17"
/locus_tag="SMA0591"
/complement(4761..4955)
/locus_tag="SMA0591"
/locus_tag="SMA0591"
/function="Miscellaneous; Hypothetical/Partial homology"
/note="glimmer prediction Excellent match to 64 amino
acids at carboxyl terminus of much larger (44 kDa) protein
from Sinorhizobium NGR234"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein, fragment"
/protein_id="AAK64967.1"
/db_xref="GI:14523392"
/translation="VVRSRWYVRKLRNFRAGIEAGISGLTRTYGLAHTWRGLHHFET
YVSSVAVYNLAFRLRPT"
/complement(5249..6421)
/locus_tag="SMA0592"
/complement(5249..6421)
/locus_tag="SMA0592"
/function="Miscellaneous; Hypothetical/Globular homology"
/note="glimmer prediction match (e-19) to two similar size
(45 kDa) proteins in Sinorhizobium sp. NGR234 over entire
length"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK64968.1"
/db_xref="GI:14523393"
/translation="WALRSANILFDEETAGTIVETANGSTRPAHSDMNEGIACCFP
STQREHEKVGGLHPPFOHLGPEGMVLEQOARSAHVVEEDGLRLRYGADCIQAVSIR
PPDDAQQLPEITEATVSPRTVSGVQKLLVTKDDENRFPVSGASALYIAKFNSSDR
IDNLVARNELSLRTAAVYGEREYVFTASLTAVIDETALIVTRDRPNEGKRLD
CAQILSKRGDYGAGYDAAYEDIAIIRHSSRAPIDLRFRNLIVFTLIGNCDH
LKNPSELTPNGTIRLSPAYDVNTATPFGGDTALISIGKEIKHLEAANOATPRAFK
EIGLPRADIDQTFKQKQVEKASITIRPPDAEPADGFRVHFKETIVDNSCRILET"
/complement(6397..7428)
/locus_tag="SMA0594"
/complement(6397..7428)
/locus_tag="SMA0594"
/function="Miscellaneous; Unknown"
/note="glimmer prediction"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK64969.1"
/db_xref="GI:14523394"
/translation="WPNNTGNTAPIAMQRYVEBARRKASGLTQKHSALAGVSHPT
MAAFERGETTLTIALALDILRVGLVDPPTBEDDQARVRDAFERWRNLVAALPDOSP
ARFPQWYRFVWTEGDLKMSBELTFEELILEAVVRKQWPFMTLPTEBALQPRVDD
LIECWLAAQGEVEVEGFNDPAPCDWRAPASGRMLIIGYODGAKETPAGIILDTTL
PLMRGQVTLIAEKIASILRQDADQAVTVHPRAMPTGIRGVLRSMANPLSLIVYEGH
GARSDEAVLEAKFSANDIESRLAECMLPLTISLYERFVAGLSLNRVAEYQRLNSP
ISKERRPRR"
/complement(7698..7919)
/locus_tag="SMA0596"
/complement(7698..7919)
/locus_tag="SMA0596"
/function="Miscellaneous; Unknown"


```

/notes="glimmer prediction"
/codon_start=1
/translate_table=11
/product="hypothetical protein"
/protein_id="AAK64970.1"
/db_xref="GI:14523395"
/translation="MCTTKITKCLFLAMQKQFASVTNNMLSLSRISCRIGLIFQINDF
DIEFATPDEGADIALAVNFKLFLKREG"

Query Match      7.9%; Score 109.6; DB 1; Length 12129;
Best Local Similarity 52.3%; Pred. No. 6,1e-11;
Matches 402; Conservative 0; Mismatches 309; Indels 57; Gaps 5;

QY 454 ATCAAGCGACATTATCTAAACCGGCTGTCGAAATAATCAAGCCGAAGCCGCGTTTC 513
DB 2997 ATGACGCGCTCGGTGCTGAAGAAGATGCTTCGCGCAAGAAAGAGCGGTGCGAGCGCC 2938
QY 514 AATATGGCCATGACCTTCGCGCGCGGATCTCAACAACATCAACCTGCTTACTGCTGGCG 573
DB 2937 AAGGTGCGCATGACCTTCGCGCGCGGATCAACGACCTCTGGAATCCGCTATTGCTGCGC 2878
QY 574 GCTGCGCGATCATCTGCTGCTACTATTGCGCGCGCGGACGATTCGCGCAATCGGC 633
DB 2877 GCGCGCGGATCAACCGCGCAAGATGTCGAACCATCTGCT----- 2834
QY 634 GCAGACGCTTGTGTGCTGACCCCGCCCGCGAAATGCGCTCAACGCTGGAAGCCGCGC 693
DB 2833 -----CCGCGCGCGGAGATGTCGCAACATGAAGTGGCG 2797
QY 694 ACCATTTTGGTTATTGCTGCGCGAGCCGCTGGAACGACGACCGCTGTTCAAGGCATA 753
DB 2796 ACCATGACCTGTTTCTGCGCTGCGCGAACCTTGGAACGACGACCTGTCACGAAAGATC 2737
QY 754 GCGCGCGCGTATCAACCGATGAAGAACTCTGAAGGACAGCGCGAAAGCTTTCGGC 813
DB 2736 GGTTCACACGCTGTCAACCGCGCGAGATTTGGCCGACGATTCGGAAGATGCTTTCGCC 2677
QY 814 GTGACCAAAACAATGCGCGAGAAATACCCCAACACTATCTGCGGTGACCAAGCGCTG 873
DB 2676 ATGCGCGCGGACTGGGTGCGAAGAACCGCGCGCTCAAGGCGCTGTCATGCGCATC 2617
QY 874 ATTGCGCGCGGATCTGCTGACGCGCGCAATTAACAAGACCGCAAGAACCATTCGAA 933
DB 2616 GAGGAAGCGCGCGAGTGTGCGGACG-----ATATGCGCAACAAAGACGAATCGCCAAG 2563
QY 934 ATGCTGCGCGCAAAACAATGCTGCTGCTGACGCTGAATGCTGCGCGCTACGATGAAC 993
DB 2562 ATGCTGCGCGCAAGCGCGAGCTGTTCAACGTGCGCGCAAGACATCTGATCGCTGAAA 2503
QY 994 GGCACCTTTCGAATACGAAAGAAAGACGATTAACGCGCGCT--ACCGGACTTCAACACCTTC 1050
DB 2502 GGTGAATACGATCAACGCGCAAGATGCTGGAAGAACGCGCGCATTCATGAATTC 2443
QY 1051 TTTCGCCACGCGCGCGACGATTCGCTCTTACAGACGTGCTGATCTGACCCAGCTG 1110
DB 2442 TGGCGGGAACACCGCTCTTACCGCTTCAAGAGCCATGACCGCTGATCTCAACGAGAAAT 2383
QY 1111 AGGCGCTGCGGCGCATGATCAATGAATCAACCGGACAACTGCTATCTGATACCGCAAG 1170
DB 2382 ATCCGCTGCGGCGAAGCT--TGCAATCGGAATACGACATCAAGGCGCTGATCGCC----- 2332
QY 1171 AACGTACCGCGCGGACATCTATCTGCGCGCGCGCAAGGAACTGCTG 1218
DB 2331 AAGGTCAACCGTACGACATCTGCGCGCAAGCGCAAGAACGACCTCGCG 2284

```

```

RESULT 15
LOCUS      AP006575      299750 bp      DNA      linear      BCT 18-OCT-2003
DEFINITION Gloeobacter violaceus PCC 7421 DNA, complete genome, section 8/16.
ACCESSION  AP006575 BA000045
VERSION     AP006575.1 GI:35212526
KEYWORDS

```

```

SOURCE      Gloeobacter violaceus PCC 7421
ORGANISM    Gloeobacter violaceus PCC 7421
REFERENCE   1
AUTHORS     Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H.,
            Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y.,
            Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N.,
            Shimo, S., Takeuchi, C., Yamada, M., and Tabata, S.
            Complete genome structure of Gloeobacter violaceus PCC 7421, a
            cyanobacterium that lacks thylakoids
            DNA Res. 10, 137-145 (2003)
TITLE       2
AUTHORS     Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H.,
            Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y.,
            Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N.,
            Shimo, S., Takeuchi, C., Yamada, M., and Tabata, S.
            Complete genome structure of Gloeobacter violaceus PCC 7421, a
            cyanobacterium that lacks thylakoids (supplement)
            DNA Res. 10, 181-201 (2003)
REFERENCE   3
AUTHORS     Kaneko, T.
            Direct Submission
            Submitted (15-AUG-2003) Takakazu Kaneko, Kazusa DNA Research
            Institute, The First Laboratory for Plant Gene Research, 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail: kaneko@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/,
            Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
            Location/Qualifiers
            1. 299750
               /organism="Gloeobacter violaceus PCC 7421"
               /mol_type="genomic DNA"
               /strain="PCC 7421"
               /db_xref="taxon:251221"
               /complement(199..606)
               /gene="g111960"
               /complement(199..606)
               /note="ORF ID:g111960
               hypothetical protein"
               /codon_start=1
               /translate_table=11
               /protein_id="BAC89901.1"
               /db_xref="GI:35212527"
               /translation="MAQKRSBALNRVEIYCPGCVLRMPGKHFFVSTTVOAREL
               LGITTKTATLHAATTTIVDRTWIRIEPFPDNGAAMIACHRKSGQVSQVATGEGW
               QQTGTVDPRRRPVSVEFTYRRSRIRLRTHG"
               /complement(797..1249)
               /gene="g111961"
               /complement(797..1249)
               /note="ORF ID:g111961
               hypothetical protein"
               /codon_start=1
               /translate_table=11
               /protein_id="BAC89902.1"
               /db_xref="GI:35212528"
               /translation="MKTVTLLIPVYTHIDFVGHVSNIYIOMEIIGRMKMEIGMP
               VHKLSQGIAPULIRITRIRYKSLKGDCAHHEIWIDELKSVSNVAAPRMKNGEGLIV
               ATDHQGVFVDVQKMRPRLPAEMERFEPLYDEAPSETRVARSRK"
               /complement(1294..3900)
               /gene="g111962"
               /complement(1294..3900)
               /note="ORF ID:g111962
               unknown protein"
               /codon_start=1
               /translate_table=11
               /protein_id="BAC89903.1"
               /db_xref="GI:35212529"
               /translation="MTHFSGQKRRRTMIFAVASVPCIMAAQVLAIPVPIPAQOTV
               AAPRVQQLGLMKGSASDLTGDSLEPMDIAPAPVQTHLQADPPQSPADTQTSQAPTA
               AAPASGDIIDVSVATRRPRTRRDTATATYTYVQDFRLALITVDALVQVFGFS
               LPALAGVNRSGNSYLRGFDORFOLLRDGISLQSRSSNNRVDISRLIEDIIVERVEVVTG

```



```
QY      924 AGCCATCGAAATGCTGGCGCAAAAACATACGTGCTGTGACGTGGAAGTGTGGCGGC 983
Db      142761 GATCCGCGGCTCTTTGCGCGACGAGTACTGCGGCGCGATCCGCTTACGTCCAGTT 142820
QY      984 TAGCATGAACGGCACTTTGATATACGAAAAAGACGATAACGGCGGCTACCGACTTCAA 1043
Db      142821 GGGCGAAGGGCGAAACCGAACACTGTGCTCTGGCAAGACCGCTCCAGTACGCCCA 142880
QY      1044 CACCTTCTTTGCGCACGGCGCCAGCTATCCGTCCTACAGCAGTGCAGTGTATCTGAC 1103
Db      142881 CCGTTTGGCGGTGCGCGGTGAACCGCCCAAGCGCAACGAGCACTGTGATCTTCAAC 142940
QY      1104 CCACTGAGGGCGCTGGGG 1121
Db      142941 CCAGATGGCTCGCTGGGG 142958
```

Search completed: May 20, 2005, 21:55:48
Job time : 6184 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 20, 2005, 05:34:38 ; Search time 811 Seconds
(without alignments)
10160.641 Million cell updates/sec

Title: US-10-689-200-1
Perfect score: 1392
Sequence: 1 atgaacaccatcatgatac.....ccggcggaaggtcgtgat 1392

Scoring table: IDENTITY NUC
Gapop 10_0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20018:*
6: geneseqn20028:*
7: geneseqn20028:*
8: geneseqn20038:*
9: geneseqn20038:*
10: geneseqn20038:*
11: geneseqn20038:*
12: geneseqn20048:*
13: geneseqn20048:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392	100.0	1392	12	ADO06542 Methylo
2	143	10.3	551	12	ADO06544 Methylo
3	85.4	6.1	747	11	ABD14925 Pseudom
4	85.4	6.1	750	11	ABD15326 Pseudom
5	82.4	5.9	1371	11	ABD15218 Pseudom
6	82.4	5.9	1371	11	ACH97397 Klebsiell
7	71.8	5.2	1209	8	ACA45921
8	69	5.0	594	11	ABD15217 Pseudom
9	69	5.0	3720	11	ABD15023 Pseudom
10	62.2	4.5	1173	8	ACA24128 Prokaryot
11	59.2	4.3	576	8	ACA23622 Prokaryot
12	58	4.2	1200	8	ACA23316 Prokaryot
13	58	4.2	1200	13	ADSI4594 Pseudom
14	58	4.2	1260	11	ABD10532 Pseudom
15	57	4.1	1116	8	ACA44309 Prokaryot
16	50	3.6	1011	9	ADA32159 DNA encod
17	49.4	3.5	1512	13	ADP43829 Bacteri
18	48.4	3.5	773	8	ACA20473 Prokaryot
19	48	3.4	744	13	ADT44805 Bacteri
20	47.8	3.4	2000	8	ADA71938 Rice gene

21	47.8	3.4	7656	8	ACA37618 Prokaryot
22	47.8	3.4	28804	2	AAT37329 Sphingom
23	47.8	3.4	28804	2	AAT92474 Sphingom
24	47.8	3.4	28804	2	AAV99812 Sphingom
25	47.8	3.4	28804	2	AAV9812 Chromosom
26	46.2	3.3	1044	12	ADJ35031 DNA encod
27	46.2	3.3	37507	12	ADH48030 Clome FS3
28	45.6	3.3	1173	8	ACA25903
29	45	3.2	8298	4	AAK72613
30	44.8	3.2	1323	12	ADM99142 Environm
31	43.4	3.1	420	11	ABD15254 Pseudom
32	43.4	3.1	531	11	ABD15367 Pseudom
33	43.4	3.1	1119	11	ABD15153 Pseudom
34	43.4	3.1	1188	8	ACA25753 Prokaryot
35	43.4	3.1	1389	11	ABD14772 Pseudom
36	43.4	3.1	5577	13	ADP44774 Bacteri
37	43	3.1	1305	8	ACA23608 Prokaryot
38	43	3.1	1308	11	ABD03949 Pseudom
39	43	3.1	1359	11	ABD03676 Pseudom
40	43	3.1	1443	11	ABD04027 Pseudom
41	43	3.1	1818	13	ADS48801 Bacteri
42	42.8	3.1	516	11	ABD10614 Pseudom
43	42.8	3.1	651	11	ACH94942 Klebsiell
44	42.6	3.1	1149	11	ABD03368 Pseudom
45	42.6	3.1	1587	11	ABD03299 Pseudom

ALIGNMENTS

RESULT 1	ADO06542	ADO06542 standard; DNA; 1392 BP.
ID	ADO06542	
XX	ADO06542;	
AC	ADO06542;	
XX	29-JUL-2004 (first entry)	
XX		
DE	Methylomonas sp 16a nitrogen transporter ntrA gene.	
XX		
KW	de; gene; C1 metabolizing bacteria; gene expression regulation;	
KW	nitrogen transporter; ntrA.	
XX		
OS	Methylomonas sp.; 16a.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1392
FT		/*tag= a
FT		/product= "nitrogen transporter ntrA"
FT		/partial
FT		/note= "no stop codon"
XX		
PN	WO2004037998-A2.	
XX		
PD	06-MAY-2004.	
XX		
XX	21-OCT-2003; 2003MO-US033698.	
PF		
XX	21-OCT-2002; 2002US-0419872P.	
PR		
XX	(DUPO) DU PONT DE NEMOURS & CO E I.	
PA		
XX	Dicoimo DJ, Picataggio SK, Seip JB, Ye RW, Wang T, Ni H;	
PI	WPI, 2004-375911/35.	
XX	P-PSDB; ADO06543.	
DR		
XX	Promoter regions responsive to presence of nitrate, acidic pH, elevated	
PT	temperatures or highly expressed in presence of methane or methanol	
PT	useful for driving expression (especially for producing zeaxanthin) in C1	
PT	metabolizing bacteria.	
XX		
PS	Claim 16; Page 69-70; 83pp; English.	

XX The present invention relates to a method of expressing a coding region
 CC of interest in a CI metabolizing bacteria. Also provided are protein,
 CC coding and promoter sequences from *Methylobacterium* sp. 16a which respond to
 CC various metabolic and growth conditions, which can be used in the method
 CC of the invention. The promoters are useful for driving expression of a
 CC number of coding regions (especially for producing zeaxanthin) in CI
 CC metabolizing bacteria under controlled conditions. The present sequence
 CC is a *Methylobacterium* sp. coding sequence.

XX Sequence 1392 BP; 347 A; 381 C; 405 G; 259 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 1392; DB 12; Length 1392;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAACCATCTTATGATCGAGCTGAAAGAAATTTGATTTGATCGCTTCTGCTA 60
 Db 1 ATGAAAACCATCTTATGATCGAGCTGAAAGAAATTTGATTTGATCGCTTCTGCTA 60
 QY 61 GCCGTTGGGGTTTGAACGATTGCGCCGATGTCGGGCGAGTTGGCAAGCTGGAAAAGAA 120
 Db 61 GCCGTTGGGGTTTGAACGATTGCGCCGATGTCGGGCGAGTTGGCAAGCTGGAAAAGAA 120
 QY 121 GATTTGAAATTCGGCTTCAATCAAGCTCACCGACATGGCGCGCTGGCGGGCGCGAA 180
 Db 121 GATTTGAAATTCGGCTTCAATCAAGCTCACCGACATGGCGCGCTGGCGGGCGCGAA 180
 QY 181 AAAGGCTTCTTGGAGAGACGAGGCGCTTCTGTCGCACTGGAACGCGACCACTGGAAG 240
 Db 181 AAAGGCTTCTTGGAGAGACGAGGCGCTTCTGTCGCACTGGAACGCGACCACTGGAAG 240
 QY 241 GTGGTATGATGATGGGTCTGTAATGGCGAACTGGACGCGCTCGCAATGCTGCGCGCGG 300
 Db 241 GTGGTATGATGATGGGTCTGTAATGGCGAACTGGACGCGCTCGCAATGCTGCGCGCGG 300
 QY 301 CGGTTAGCGGCGAGGCTTGGCTTCGCGACACGAGGCGGATATCGAGGTGGCGTTCAAGATG 360
 Db 301 CGGTTAGCGGCGAGGCTTGGCTTCGCGACACGAGGCGGATATCGAGGTGGCGTTCAAGATG 360
 QY 361 GCGCTTCACGCGAACCGCATTAACGATGTCGATCAATGAATCTGGATCAGATGAAGCCGAAC 420
 Db 361 GCGCTTCACGCGAACCGCATTAACGATGTCGATCAATGAATCTGGATCAGATGAAGCCGAAC 420
 QY 421 ATACCGCTGGAAGGCGGTAAACGGGTGATCCGATCAAGGCGATTAATCTAAACCGGTC 480
 Db 421 ATACCGCTGGAAGGCGGTAAACGGGTGATCCGATCAAGGCGATTAATCTAAACCGGTC 480
 QY 481 GTGGAATAATACAAAGCCGAAAGCGGTTCAATATGCGATGAACCTTCCCGGCGGA 540
 Db 481 GTGGAATAATACAAAGCCGAAAGCGGTTCAATATGCGATGAACCTTCCCGGCGGA 540
 QY 541 TCTCAACATCAAACTGCGTTACTGCTGCGGCTGCGGCTATCAATCTTGGCTACTAT 600
 Db 541 TCTCAACATCAAACTGCGTTACTGCTGCGGCTGCGGCTATCAATCTTGGCTACTAT 600
 QY 601 TGGCGGCGCGAGGACATTTCCGGCCAAATCGGCGCAGACGCTTGTGTGCGTGACCCG 660
 Db 601 TGGCGGCGCGAGGACATTTCCGGCCAAATCGGCGCAGACGCTTGTGTGCGTGACCCG 660
 QY 661 CGGCGGCAAAATGCGGTCCACGCGTGAAGGCGGACCAATTTGGTTATGCGTGGCGAG 720
 Db 661 CGGCGGCAAAATGCGGTCCACGCGTGAAGGCGGACCAATTTGGTTATGCGTGGCGAG 720
 QY 721 CCGTGAACACAGAGCGGTGTTCAAGGGCATAGCGTCCGGTGATCAACCGTGAAGAA 780
 Db 721 CCGTGAACACAGAGCGGTGTTCAAGGGCATAGCGTCCGGTGATCAACCGTGAAGAA 780
 QY 781 CTCTGGAAGGACACGCGGAAAAAGTCTTCCGCGTGACCAAAATGCGGCGGAAATAC 840
 Db 781 CTCTGGAAGGACACGCGGAAAAAGTCTTCCGCGTGACCAAAATGCGGCGGAAATAC 840
 QY 841 CCCAACACCTATCTGCGGTGACCAAGGCGCTGATTCGGGCGCGCATCTGGCTGAGCGCC 900

Db 841 CCCAACACCTATCTGCGGTGACCAAGGCGCTGATTCGGGCGCGCATCTGGCTGAGCGCC 900
 QY 901 GACATTAACAGAACCCGAAAGAGCCATGAAATCTGCGCGCAAAAACATACGTGGT 960
 Db 901 GACATTAACAGAACCCGAAAGAGCCATGAAATCTGCGCGCAAAAACATACGTGGT 960
 QY 961 GCTGACGTGAAGTGTGGCGGTGACATGAACGCGCATCTTTGAAATCGAAAAAGCGAT 1020
 Db 961 GCTGACGTGAAGTGTGGCGGTGACATGAACGCGCATCTTTGAAATCGAAAAAGCGAT 1020
 QY 1021 AAACGCGGCTACCGGACTTCAACACCTTCTTCCGACAGGCGCGCATTCGGTCTAC 1080
 Db 1021 AAACGCGGCTACCGGACTTCAACACCTTCTTCCGACAGGCGCGCATTCGGTCTAC 1080
 QY 1081 AGCAGTGAAGTGTGATGACCCAGCTGAGGCGCTGGGCGCATGATCAATGAATCAAA 1140
 Db 1081 AGCAGTGAAGTGTGATGACCCAGCTGAGGCGCTGGGCGCATGATCAATGAATCAAA 1140
 QY 1141 CCGGACACTGATATCTGATACCGCAAGAACGTCTACCGCCGGAATCTATCTCGCC 1200
 Db 1141 CCGGACACTGATATCTGATACCGCAAGAACGTCTACCGCCGGAATCTATCTCGCC 1200
 QY 1201 GCGGCGCAAGAACTGGTCCGACAGAGGCAAGGCGCAAGGCGCAAGACTTCCCTGCGATACC 1260
 Db 1201 GCGGCGCAAGAACTGGTCCGACAGAGGCAAGGCGCAAGGCGCAAGACTTCCCTGCGATACC 1260
 QY 1261 TCGATCAACCGCTCGCAGAAATTTCTTCATCGACAAAGTCCGTTGATGCCAACAGGCC 1320
 Db 1261 TCGATCAACCGCTCGCAGAAATTTCTTCATCGACAAAGTCCGTTGATGCCAACAGGCC 1320
 QY 1321 AACGATTACTCGCCCAAGTTTGGATAGGTCTGAAAGGCAAGCAACCGTAGCCGCGGC 1380
 Db 1321 AACGATTACTCGCCCAAGTTTGGATAGGTCTGAAAGGCAAGCAACCGTAGCCGCGGC 1380
 QY 1381 AAGGTGATGAT 1392
 Db 1381 AAGGTGATGAT 1392

RESULT 2
 ADO06544
 ID ADO06544 standard; DNA; 551 BP.
 XX ADO06544;
 AC
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE *Methylobacterium* sp 16a nitrogen transporter nrtA gene promoter.
 XX
 KW ds; gene; CI metabolizing bacteria; gene expression regulation;
 KW nitrogen transporter; nrtA; promoter.
 XX
 OS *Methylobacterium* sp.; 16a.
 XX
 PN WO2004037998-A2.
 XX
 PD 06-MAY-2004.
 XX
 PF 21-OCT-2003; 2003WO-US033698.
 XX
 PR 21-OCT-2002; 2002US-0419872P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Dicosimo DJ, Picetaggio SK, Seip JE, Ye RW, Wang T, Ni H;
 XX WPI; 2004-375911/35.
 XX
 PT Promoter regions responsive to presence of nitrate, acidic pH, elevated
 PT temperatures or highly expressed in presence of methane or methanol
 PT useful for driving expression (especially for producing zeaxanthin) in CI
 PT metabolizing bacteria.

XX Claim 37, Page 72, 83pp; English.
PS
XX The present invention relates to a method of expressing a coding region
CC of interest in a CI metabolizing bacteria. Also provided are protein,
CC coding and promoter sequences from *Methylobacterium* sp. 16a which respond to
CC various metabolic and growth conditions, which can be used in the method
CC of the invention. The promoters are useful for driving expression of a
CC number of coding regions (especially for producing zeaxanthin) in CI
CC metabolizing bacteria under controlled conditions. The present sequence
CC is a *Methylobacterium* sp. promoter sequence.
XX
SQ Sequence 551 BP; 139 A; 119 C; 135 G; 158 T; 0 U; 0 Other;
Query Match 10.3%; Score 143; DB 12; Length 551;
Best Local Similarity 100.0%; Pred. No. 7e-27;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAACCATCTAGATGAGCTCGAAGAAATTGTTATGACGTTATCGCTTGGCTA 60
DB 409 ATGAAACCATCTAGATGAGCTCGAAGAAATTGTTATGACGTTATCGCTTGGCTA 468
QY 61 GCGGTTGGGCTTGAAGATTTGGCCGAGATGCGGGCAGTTGGCAAGCTGAAAGGAA 120
DB 469 GCGGTTGGGCTTGAAGATTTGGCCGAGATGCGGGCAGTTGGCAAGCTGAAAGGAA 528
QY 121 GATTGAATTCGGCTTCATCAA 143
DB 529 GATTGAATTCGGCTTCATCAA 551
RESULT 3
ABD14925/c
ID ABD14925 standard; DNA: 747 BP.
AC ABD14925;
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polynucleotide #13529.
DE
XX
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KM antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI WPI; 2003-615309/58.
DR P-PSDB; ABO81354.
XX
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX
XX Disclosure; SEQ ID NO 13529; 455bp; English.
PS
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC

CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABD01397-
CC ABD1967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 747 BP; 114 A; 234 C; 281 G; 118 T; 0 U; 0 Other;
Query Match 6.1%; Score 85.4; DB 11; Length 747;
Best Local Similarity 51.3%; Pred. No. 7.5e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;
QY 651 GGTGACCCCGCCCGCCCAATGCGTTCACGCTGGAAGCCGCAACATTTCCGTTATG 710
DB 559 GGTGATTTCCACCGCCGAGATGCTGCGCATCTCCAGGCCGCGCATCGACGGGTTCTG 500
QY 711 CGTGGCGAGCGGTGGAACAGCAGGCGGTTCAAGGCGATAGGCGTGGCGGTATCAC 770
DB 499 CGCGGGGCGCCCTTGGGGCGCCCTTGGCGTGAACAGGGCAGGGCTTCACTCGCCAC 440
QY 771 CGATGAAGAACTGTGAAGAGCAGCCGCGAAGAAAGTCTTGGCGGTGACCAACATGGGC 830
DB 439 CAGCCAGGCGATTTGGCCGCGAACAACCGGAAAGTCTTGGCGATCAACCGCGCTTCTG 380
QY 831 GGAAGAAATACCCCAACCTATTTGGCGGTGACCAAGCGGTGATTTGGCGCGGATCTG 890
DB 379 CGAGGCTTACCCGAAACACCGCCGCGGTGCTATGGC--GATCTCGACGCCAGCCG 323
QY 891 GCTGAGCGCCGAGCAATTAACAAGAACCGCAAGCAAGCATGAAATGCTGGCGGAAAAACA 950
DB 322 CTTCATCGAACAAGAACCGCAAGCAAGCATGAAATGCTGGCGGAAAAACA 263
QY 951 ATAGCTCGGTGCTGAGTGAAGTGTGGCGGTAGCATGAAGCGCACTTTCGAATACGA 1010
DB 262 CTATGTGAGCTTCGCTCGCGCGCATCCAGCGCGTCTTGGCGGTACCAAGACGG 203
QY 1011 AAAAGACGATMAACGCGCGTACCGGACCTTCAACACTTCTTTCGCCAGCGGCCAGCTA 1070
DB 202 CTCGGCAAGCGCTGGAGAGACCAACATCCGCTGCTTCTTACGCCAGCGCAAGTCAA 143
QY 1071 TCCGTCT--ACAGCAGTCACTGTGTATCTGACCCAGCTAGAGCGCTGGCGATGAT 1127
DB 142 CGGCGCTGAGCTTCGATGCGATGTGTTCATGACCACTTCGCGCTGGCGCTGCT 83
QY 1128 CAATGAA 1134
DB 82 GCGCGAA 76
RESULT 4
ABD15326
ID ABD15326 standard; DNA: 750 BP.
AC ABD15326;
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polynucleotide #13930.
DE
XX
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KM antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
PN 22-APR-2003.
XX

XX 18-FEB-1999; 99US-00252991.
PF 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
PI P-PSDB; ABO81755.
DR MPI; 2003-615309/58.
XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 13930; 455pp; English.

CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences ABD01397-
CC ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX

Sequence 750 BP; 115 A; 293 C; 237 G; 105 T; 0 U; 0 Other;

Query Match 6.1%; Score 85.4; DB 11; Length 750;
Best Local Similarity 51.3%; Pred. No. 7.5e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

```
QY 651 GGTGACCCCGCCGCGCAATGCGTCCAGTGGAGGACGACCATTTGGTTATG 710
    |||||
DB 54 GGTGTTCCACCGCGCGCAATGTCGCATCTCCAGCGCGCATGACGGTTCTG 113
QY 711 CGTCGCGAGCGGTGGAACGACGAGCCGCTTCAAGCATAGCGTCCGTTATC 770
    |||||
DB 114 CCGCGCGCGCCCTGCGGGGCGCTGCGCTGACGACGAGGCGCTTACCATGCCAC 173
    |||||
QY 771 CGATGAGAACTCTTGGAAGACACGCGGAAAAAGTCTTGCGCGTGAACCAATGCGC 830
    |||||
DB 174 CAGCCGAGCGATCTGCGCGACACCGGAAAAAGTCTTGCGCATACCGCGCTTCGT 233
    |||||
QY 831 GGAGAAATCCCAACCTATCTGCGCGTGAACGAGCGCGTATTCGCGCGCATGTCG 890
    |||||
DB 234 CCAACGCTACCGGAACCGCGCGCGTGTGATGAGC--GCTCTGACCGCCAGCGG 290
    |||||
QY 891 GCTGACGCGCGCAATTAACAAGAACGCAAGAGGCATCGAAATCTGCGCGCAAAAACA 950
    |||||
DB 291 CTTTCATCGAAGACGCGCGAAGACCGCTGCGACCGCGCAATCATACGCGCGCGCA 350
    |||||
QY 951 ATAGCTGGTGTGACGTGGAAGTGTGCGCGCTAGCATGAACGGGACATTTCGAATACGA 1010
    |||||
DB 351 CTATGTCGACGCTCGCTCGCGCGCATTCAGCGCGCGCTTCTTGCGCGCGCAAGTCAA 410
    |||||
QY 1011 AAAAGACGTAACGCGCGCTACCGGACATTCTTTCCGACGCGCGCGCAAGCTTA 1070
    |||||
DB 411 CCTCGGCAACGCGCTGCGACGACCATCCGCTGCGTTCTTACGCGGCGCGCAAGTCAA 470
    |||||
QY 1071 TCCGCTCT--ACAGAGTGTGAGTGTATATGACCGACGCTGAGGCGGTGGGGCATGAT 1127
    |||||
DB 471 CCGGCGCTGCTCTCGGATGGCATGTGTTTCATGACCGCATGTTCCGCGCTGGGGCTGCT 530
    |||||
```

QY 1128 CATGAA 1134
DB 531 GCGCGAA 537

RESULT 5
ABD15218
ID ABD15218 standard; DNA; 831 BP.
XX
XX ABD15218;
AC 29-JUL-2004 (first entry)
XX
XX
XX

DE *Pseudomonas aeruginosa* polynucleotide #11822.
XX
XX Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection;
XX antibacterial.
XX

OS *Pseudomonas aeruginosa*.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, DeLoughery C, Bush D;

DR P-PSDB; ABO81647.

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX
XX Disclosure; SEQ ID NO 13822; 455pp; English.

CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences ABD01397-
CC ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX

Sequence 831 BP; 124 A; 327 C; 259 G; 121 T; 0 U; 0 Other;

Query Match 6.1%; Score 85.4; DB 11; Length 831;
Best Local Similarity 51.3%; Pred. No. 7.8e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

```
QY 651 GGTGACCCCGCCGCGCAATGCGTCCAGTGGAGGACGACCATTTGGTTATG 710
    |||||
DB 100 GGTGTTCCACCGCGCGCAATGTCGCATCTCCAGCGCGCATGACGGTTCTG 159
    |||||
QY 711 CGTCGCGAGCGGTGGAACGACGAGCCGCTTCAAGCATAGCGTCCGTTATC 770
    |||||
DB 160 CCGCGCGCGCCCTGCGGGGCGCTGCGCTGACGACGAGGCGCATACCATGCCAC 219
    |||||
```


CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ. Sequence 1209 BP; 229 A; 395 C; 376 G; 209 T; 0 U; 0 Other;

Query Match 5.2%; Score 71.8; DB 8; Length 1209;

Best Local Similarity 53.8%; Pred. No. 3e-08; Mismatches 127; Indels 0; Gaps 0;

Matches 188; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
QY 102 TGGCAAGCTGGAAAGAAAGATTGAATTGGCTTCATCAAGCTGACCGACATGGCGCC 161
DB 45 TGAACGCCCGCGAATGAACAGCTGAGACGTGGGCTTCATGACGACGACGTGGCGCC 104
QY 162 GCTGGGGGTGGCCCGCGAAGAAAGCTTCTTCAGAGACGAGGCTGTTCGTAATCGA 221
DB 105 GCTGGGTGGTGGTCCACCCAGGGGTTTGGCCAGCTTACGGGCTGAGCTGAACCTGAA 164
QY 222 AGCGCAGGCAATGGAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 281
DB 165 AGCCGAGACTCTCTGGGCGGCGGCTGCGTACAGAGCTGTGACGGGAGCTTTAGGCGGC 224
QY 282 GCAATGCTGGCGCGCGCGCTTACCGGCAAGCTTGGCTTGGGACCAAGGCCGATAT 341
DB 225 ACATAGCTGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 284
QY 342 CGAGGTGCGCTTACGATGGGCTTCAACGGCAACG 376
DB 285 CGACATGGCCATCTCATGGGCTCAACGAAACG 319

RESULT 8

ABD15217 standard; DNA; 594 BP.

ABD15217;

29-JUL-2004 (first entry)

Pseudomonas aeruginosa polymnucleotide #13821.

Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
antibacterial.

Pseudomonas aeruginosa.

US6511795-B1.

22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR MPI; 2003-615309/58.

DR P-PSDB; ABO81646.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 13821; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biophysics technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX SQ. Sequence 594 BP; 102 A; 208 C; 196 G; 88 T; 0 U; 0 Other;

Query Match 5.0%; Score 69; DB 11; Length 594;

Best Local Similarity 52.2%; Pred. No. 1.3e-07; Mismatches 140; Indels 0; Gaps 0;

Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 84 GCCCGATGTCGGGGGCGGCTTGGCAAGCTGGAAAGAAAGATTGAATTGGGCTTCATCA 143
DB 204 GGGCTGGGCGCGCGGCTTCGACGACCGAAGAAAGCGCTTGACATCGTTTCATGGC 263
QY 144 GCTCACCGCATGTCGCGCGCTGCGGTGGCGCGCCGCAAAAAGCTTCTTTCAGAGACGAGG 203
DB 264 GCTGACCGAATTCGCGCTGCTGATGCTGCGCCGACCCAGGCTTGGCCAGCCCTACG 323
QY 204 CCGTTCGTCGCACTGGAAGCGGCAAGCTGGAAGGTGTGATGATGATGATGATGATGATGAT 263
DB 324 CCGTACGCTCAATCTTCAGGCGGCAACCGTCTGGGCGACCTTGGCGGACAACTGCTCAG 383
QY 264 TGGCGAATGGAAGGCTGCGCATGCTGCGCGCGGCGCGCTTACGCGGCGGCTTGGCTT 323
DB 384 CGCGAGCTGGAATGCGCGGCGGCAAGCTGTGACGGGCTTGTCTACGCGGCTTCACTCGGCTT 443
QY 324 CGGCAACGAAGCGGATGATGAGGTGCGCTTTCAGATGGGCTTCAACGCGGCAACG 376
DB 444 CGGCGGCGGCGGCGGCAAGGATGCGGTGCTCATGGGACGTGTGCCAAGAG 496

RESULT 9

ABD15023/c standard; DNA; 3720 BP.

ABD15023;

29-JUL-2004 (first entry)

Pseudomonas aeruginosa polymnucleotide #13627.

Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;

XX antimicrobial.
XX Pseudomonas aeruginosa.
OS US6551795-B1.
PN 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PI Rubenfield MJ, Nollig J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
DR P-PSDB; AB081452.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
PS Disclosure; SEQ ID NO 13627; 455bp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD1967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 3720 BP; 576 A; 1147 C; 1326 G; 671 T; 0 U; 0 Other;
Query Match 5.0%; Score 69; DB 11; Length 3720;
Best Local Similarity 52.2%; Pred. No. 2.4e-07;
Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 84 GCCCGATGTCGGGCGCAAGTTGGCAAGTGAAGAAAGATTGCGGCTTCATCAA 143
DB 676 GCGCTGGGCGCGGCTTCGACGACCGGAAAGAGCGCCCTGACATCGGTTTCATGCG 617
QY 144 GCTCACCGACATGCGCGCTGGCGGTGCGCGCCGAAAGAGCTTCTTCAGACGAGGG 203
DB 616 GCTGACCGAATCGGCTCGCTGATCGTGGCGCCACCAAGGCGCTTCCGCAAGCCCTACGG 557
QY 204 CCGTTCGTGCACTGGAAGCGGCAAGTGAAGTGTATGATGATGATGAGTGGTGGAA 263
DB 556 CTTGACGCTCAATCTCAGGCGCCAAACGCTCTGGGCCACCTTGGCGCAAGAGCTGCAG 497
QY 264 TGGCGAAGTGAAGCGCTTCGACATGCTGCGCGCGCGCCGTTAGCGGCCAGCGTTGGCTT 323
DB 496 CCGCGAGCTGATGCGCGCGGACGAGTCCGTGACGCGCTTGTACGCGCTCCAGCTCGGCTT 437
QY 324 CCGGACCAAGCGCGGATTCGAGAGTGCCTTTCAGATGCGGCTTCACGCGGCAAG 376
DB 436 CCGCGGCGAGCGCGGCAAGGATGCGCGTGTCTGATGAGTGTGCAAAACG 384

RESULT 10
ACA24128
ID ACA24128 standard; DNA; 1173 BP.

XX ACA24128;
AC 19-JUN-2003 (first entry)
XX
DT Prokaryotic essential gene #5785.
XX
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Borrelia cepacia.
XX W0200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX P-PSDB; AB020258.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 11998; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; (7) identifying a gene in an operon required for
CC proliferation; (8) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (9)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism's activity; (10) manufacturing an antibiotic; (11) profiling a
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1173 BP; 180 A; 396 C; 404 G; 193 T; 0 U; 0 Other;
Query Match 4.5%; Score 62.2; DB 8; Length 1173;

Best Local Similarity 58.4%; Pred. No. 9.4e-06;
Matches 128; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 131 TCGGCTTCATCAAGCTCACCAGACATGCGCGCTGGCGTGGCCGCGAAGAGCTTCT 190
DB 116 TCGGCTATCTGCAGATCACCAGACGCGGCCCGCTGCTGGTGGCCCAACAAAGGCTATT 175
QY 191 TCGAGACGAGGGCCCTGT---TCGTCAACTGGAAGCGCAGGCCCAACTGGAAGTGTGA 247
DB 176 TCGCGTCCGAAAGGCTGACGCGTGAACCAACCGAAGCTGCTGCTACTGGGCGCAACTCG 235
QY 248 TGGATATGGGTTCGTAATGAGCACTGAGCGGCTCGACATGTGTGGCGCGCGCTTAG 307
DB 236 TCGAGGCGTTCTCTGTCCGCGCAGGTCAAGCTGTGCACTGCTGTGGCCGATACCTGT 295
QY 308 CGGCCAGCGTTGGCTTTCGGACCAAGCGCGATATCGAG 346
DB 296 GGGCGCGTACGCGACCGCGCGCGCGCGAAGGTGTGG 334

RESULT 11

ACA23622
ID ACA23622 standard; DNA; 576 BP.

ACA23622;

19-JUN-2003 (first entry)

Prokaryotic essential gene #5279.

Antisense; ds: prokaryotic essential gene; cell proliferation;
drug design; gene.

Borrelia cepacia.

W020027183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342823P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykind JW,
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
P-PSDB; ABU19752.

WPI; 2003-029926/02.

New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 11492; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 576 BP; 84 A; 206 C; 194 G; 92 T; 0 U; 0 Other;

Query Match 4.3%; Score 59.2; DB 8; Length 576;

Best Local Similarity 57.9%; Pred. No. 4.5e-05;
Matches 125; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 131 TCGGCTTCATCAAGCTCACCAGACATGCGCGCTGGCGTGGCCGCGAAGAGCTTCT 190
DB 116 TCGGCTATCTGCAGATCACCAGACGCGGCCCGCTGCTGGTGGCCCAACAAAGGCTATT 175
QY 191 TCGAGACGAGGGCCCTGT---TCGTCAACTGGAAGCGCAGGCCCAACTGGAAGTGTGA 247
DB 176 TCGCGTCCGAAAGGCTGACGCGTGAACCAACCGAAGCTGCTGCTACTGGGCGCAACTCG 235
QY 248 TGGATATGGGTTCGTAATGAGCACTGAGCGGCTCGACATGTGTGGCGCGCGCTTAG 307
DB 236 TCGAGGCGTTCTCTGTCCGCGCAGGTCAAGCTGTGCACTGCTGTGGCCGATACCTGT 295
QY 308 CGGCCAGCGTTGGCTTTCGGACCAAGCGCGATATCGAG 343
DB 296 GGGCGCGTACGCGACCGCGCGCGCGCGCGAAGGTGG 331

RESULT 12

ACA42316
ID ACA42316 standard; DNA; 1200 BP.

ACA42316;

19-JUN-2003 (first entry)

Prokaryotic essential gene #23973.

Antisense; ds: prokaryotic essential gene; cell proliferation;
drug design; gene.

Pseudomonas aeruginosa.

W020027183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342823P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU38446.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 30186; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1200 BP; 190 A; 420 C; 409 G; 181 T; 0 U; 0 Other;
Query Match 4.2%; Score 58; DB 8; Length 1200;
Best Local Similarity 56.2%; Pred. No. 0.00012;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;
QY 131 TCGGCTTCATCAAGCTCAGCAGCATGAGCGCGCTGGGGGCGCCGGAAGAGCTTCT 190
DB 137 TCGGCTTACTGCGCATATCCAGCCAGCCCTTGTGCTGAGCCATGCCAGCTGT 196
QY 191 TCGAGGACGAGGAGCTGTGCTGTCGCACTGAGAGCGCAGCC--AACTGAGAGTGTGA 247
DB 197 TCGAGGCGGAGGAGATCAGAGCGGAACGCGCGTGTGCTGTCGATGAGTGGCCGAGTGA 256
QY 248 TGAATAGGCTGCTGAATGCGGAATGAGCGCTTCGACATGCTGCGCCCGCGCTTAG 307
DB 257 TCGAGGCGTTCATCTCCGCGCAGGTCAACGTCACTCACTGCTGTCCCGCATGACCTCT 316
QY 308 CGGCGAGCGTTGGCTTTCGGGACCAAGCGCGATATCGAGTGGCGTTGAGATGGGCTTCA 367
DB 317 GGGGCC-----CGCTACGGGACGAGGTGCCGCGCAAGGTGTGGCTTGAACCAAGTTCG 370
QY 368 ACGGCAACGCGATTAACGCTGTCCATGAATC 399
DB 371 GCGGCTCGGGGCTCAGCGTGGCTCCGAGATC 402
RESULT 13
ADSI4594
ID ADSI4594 standard; DNA; 1200 BP.

XX
AC ADSI4594;
XX
XX 16-DEC-2004 (first entry)
DT
XX
DE Pseudomonas aeruginosa quorum sensing controlled gene PA2328, SEQ ID 149.
XX
XX Cytostatic; Anti-HIV; Antimicrobial; Immunosuppressive;
XX quorum sensing signalling; bacterium; quorum sensing controlled gene;
XX biofilm-associated disease; cystic fibrosis; AIDS; middle ear infection;
XX acne; periodontal disease; gene; der; PA2328.
XX
OS Pseudomonas aeruginosa.
XX
XX WO2004083385-A2.
XX
XX 30-SEP-2004.
XX
XX 11-MAR-2004; 2004WO-US007467.
XX
XX 14-MAR-2003; 2003US-00389647.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Greenberg EP, Schuster M, Lostroh C;
XX
XX WPI; 2004-709932/69.
XX
PT Identifying a modulator of quorum sensing signaling in bacteria, useful
PT for treating a biofilm-associated disorder, comprises contacting the cell
PT with a quorum sensing signal molecule in the presence and absence of a
PT test compound.
XX
PS Disclosure; SEQ ID NO 149; 233bp; English.
XX
XX The present invention relates to a method for identifying a modulator of
XX quorum sensing signalling in bacteria. The method comprises: providing a
XX cell that comprises a quorum sensing controlled gene (ADSI4446-ADSI4798),
XX where the cell is responsive to a quorum sensing signal molecule such
XX that a detectable signal is generated; contacting the cell with a quorum
XX sensing signal molecule in the presence and absence of a test compound;
XX and detecting a change in the detectable signal. The method and modulator
XX identified by the method are useful for treating a biofilm-associated
XX disease or disorder, e.g. cystic fibrosis, AIDS, middle ear infections,
XX acne, periodontal disease, catheter-associated infections, and medical
XX device-associated infections. Note: The sequence data for this patent was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1200 BP; 190 A; 420 C; 409 G; 181 T; 0 U; 0 Other;
Query Match 4.2%; Score 58; DB 13; Length 1200;
Best Local Similarity 56.2%; Pred. No. 0.00012;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;
QY 131 TCGGCTTCATCAAGCTCAGCAGCATGAGCGCGCTGGGGGCGCCGGAAGAGCTTCT 190
DB 137 TCGGCTTACTGCGCATATCCAGCCAGCCCTTGTGCTGAGCCATGCCAGCTGT 196
QY 191 TCGAGGACGAGGAGCTGTGCTGTCGCACTGAGAGCGCAGCC--AACTGAGAGTGTGA 247
DB 197 TCGAGGCGGAGGAGATCAGAGCGGAACGCGCGTGTGCTGTCGATGAGTGGCCGAGTGA 256
QY 248 TGAATAGGCTGCTGAATGCGGAATGAGCGCTTCGACATGCTGCGCCCGCGCTTAG 307
DB 257 TCGAGGCGTTCATCTCCGCGCAGGTCAACGTCACTCACTGCTGTCCCGCATGACCTCT 316
QY 308 CGGCGAGCGTTGGCTTTCGGGACCAAGCGCGATATCGAGTGGCGTTGAGATGGGCTTCA 367
DB 317 GGGGCC-----CGCTACGGGACGAGGTGCCGCGCAAGGTGTGGCTTGAACCAAGTTCG 370
QY 368 ACGGCAACGCGATTAACGCTGTCCATGAATC 399
DB 371 GCGGCTCGGGGCTCAGCGTGGCTCCGAGATC 402

Dh 371 GCGGCTCGGGGCTCAGCGTGGCTCCGGAGATC 402

RESULT 14
ABD10532
ID ABD10532 standard; DNA; 1260 BP.
XX
XX ABD10532;
AC
XX 29-JUL-2004 (first entry)
DT
XX Pseudomonas aeruginosa polynucleotide #9136.
DE
XX
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KM antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
PN
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI; 2003-615309/58.
DR P-PSDB; ABO76961.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 9136; 455bp; English.
PS
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
SQ Sequence 1260 BP; 202 A; 450 C; 423 G; 185 T; 0 U; 0 Other;

Query Match 4.2%; Score 58; DB 11; Length 1260;
Best Local Similarity 56.2%; Pred. No. 0.00012;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

Qy 131 TCGGCTTCATCAAGCTCACCAGCATGGCCGCTGGGCGCCGGAAGAGCTTCT 190
Db 197 TCGGCTACTGCGGACACCGACCGCCCTGCTGGTGGCCATGCCAAGGCTGT 256
Qy 191 TCGAGCAGCAGGCGCTGTTGCTGCAACTGGAAGCCGAGGCC--AACTGGAAGTGTGA 247
Db 257 TCGAGGCGCAGGCGCATCCAGCGCAACGCGCGTGTGTTGGCTAGCTGGCGCAGGTGA 316
Qy 248 TGGATAGGCTCGTGAATGGCGAACTGAGCGGCTCGCACTGCTGGGCGCGCGCTTNG 307

Dh 317 TCGAGGCGTTCATCTCCGGCAGGTCAAGTCATCCACTGCTGTGCGCGAGTACCGTCT 376

Qy 308 CGGCGACCGCTTGCTTCGCGACCAAGCCGAGATTCAGGTGCCGTTACAGATGGCTTCA 367
Db 377 GGGCC-----CGCTACGCGACGACAGTCCCGCAGAGTGGTGCTGGAACCAAGTCG 430

Qy 368 ACGGCAACCGCATTAAGTGTGTCATGAATC 399
Db 431 GCGGCTCGGGGCTCAGCGTGGCTCCGAGATC 462

RESULT 15
ACA44309
ID ACA44309 standard; DNA; 1116 BP.
XX
XX ACA44309;
AC
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Prokaryotic essential gene #25966.
DE
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
XX Pseudomonas putida.
OS
XX WO200277183-A2.
PN
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI
XX WPI; 2003-029926/02.
DR P-PSDB; ABU40439.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 32179; 1766bp; English.
PS
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-regulated gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antitense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 1116 BP; 215 A; 361 C; 342 G; 198 T; 0 U; 0 Other;

Query Match 4.1%; Score 57; DB 8; Length 1116;

Best Local Similarity 54.1%; Pred. No. 0.00021;

Matches 139; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

```
QY 131 TCAGGCTTCATCAAGCTCAACGACATGAGCGCGCTGCGGTGGCCGCGAAGAGCTTCT 190
    |||||
DB 56 TCAGGCTTCATCAAGCTCAACGACATGAGCGCGCTGCGGTGGCCATPACACAGGCTGT 115
    |||||
QY 191 TCAGGAGAGAGGCT---GTTGCTGCACTGGAAGCGCAGCCCACTGGAAGGTGTGA 247
    |||||
DB 116 TCAGGAGAGAGGCTCAAGGCTCAAGGCGCGGCTGCTGCGGAGCTGGGCGCAGGTGA 175
    |||||
QY 248 TGAATAGGCTGCTGAATGCGAACTGACGCGCTGCGCATGCTGCGCGCGCGCTTAG 307
    |||||
DB 176 TCAGGAGATTCATCTCGGCGCAGGTCAATGTCATCCACTGCTTTGCGCGATGACGAT 235
    |||||
QY 308 CGGCCAGCGCTGGCTTGGGACCAAGGCGCATATCGAGTGCCTTACGATGGGCTTCA 367
    |||||
DB 236 GGGCGGCGCTATGCGACAGGAGTGCACAGGCGAAGGTGTGCGCTGGAACACAGTGGCGGTT 295
    |||||
QY 368 ACGGCAACGCGATTAAG 384
    |||||
DB 296 CGGCGCTGACTGTTGCG 312
    |||||
```

Search completed: May 20, 2005, 20:12:42
Job time : 817 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 20, 2005, 19:00:43 ; Search time 259 Seconds
(without alignments)
8794.198 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaaacacatcattagatc.....ccgcgcgaagtcgtcgat 1392

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	85.4	6.1	747	US-09-252-991A-13529	Sequence 13529, A
2	85.4	6.1	750	US-09-252-991A-13930	Sequence 13930, A
3	85.4	6.1	831	US-09-252-991A-13822	Sequence 13822, A
4	82.4	5.9	1371	US-09-489-039A-13192	Sequence 13192, Ap
5	69	5.0	594	US-09-252-991A-13821	Sequence 13821, A
6	69	5.0	3720	US-09-252-991A-13627	Sequence 13627, A
7	58	4.2	1260	US-09-252-991A-9136	Sequence 9136, Ap
8	50	3.6	1011	US-09-328-352-3446	Sequence 3446, Ap
9	48.2	3.5	1368	US-09-902-540-7805	Sequence 7805, Ap
C 10	48.2	3.5	6555	US-09-902-540-777	Sequence 777, App
11	48	3.4	750	US-09-902-540-5834	Sequence 5834, Ap
12	48	3.4	72704	US-09-902-540-1273	Sequence 1273, Ap
C 13	47.8	3.4	28804	US-08-592-874-1	Sequence 1, App11
C 14	47.8	3.4	28804	US-09-096-942-2	Sequence 2, App11
C 15	47.8	3.4	28804	US-09-096-942-2	Sequence 2, App11
C 16	44.4	3.2	7218	US-08-232-463-14	Sequence 14, App1
C 17	44.4	3.1	3384	US-09-902-540-3780	Sequence 3780, Ap
C 18	43.8	3.1	23233	US-09-902-540-1184	Sequence 1184, Ap
19	43.4	3.1	420	US-09-252-991A-13858	Sequence 13858, A
20	43.4	3.1	531	US-09-252-991A-13971	Sequence 13971, A
21	43.4	3.1	1119	US-09-252-991A-13757	Sequence 13757, A
C 22	43.4	3.1	1389	US-09-252-991A-13376	Sequence 13376, A
C 23	43.4	3.1	5523	US-09-902-540-2081	Sequence 2081, Ap
C 24	43.4	3.1	5577	US-09-902-540-5383	Sequence 5383, Ap
25	43.4	3.1	34552	US-09-902-540-1262	Sequence 1262, Ap
26	43	3.1	1308	US-09-252-991A-2553	Sequence 2553, Ap
C 27	43	3.1	1359	US-09-252-991A-2280	Sequence 2280, Ap

28	43	3.1	1443	US-09-252-991A-2631	Sequence 2631, Ap
29	42.8	3.1	516	US-09-252-991A-9218	Sequence 9218, Ap
30	42.8	3.1	651	US-09-489-039A-737	Sequence 737, App
31	42.6	3.1	1149	US-09-252-991A-1972	Sequence 1972, Ap
C 32	42.6	3.1	1587	US-09-252-991A-1903	Sequence 1903, Ap
33	42.6	3.1	1836	US-09-252-991A-2056	Sequence 2056, Ap
34	42.4	3.0	450	US-09-489-039A-2830	Sequence 2830, Ap
C 35	42.4	3.0	465	US-09-252-991A-1459	Sequence 1459, Ap
C 36	42.4	3.0	939	US-09-489-039A-2942	Sequence 2942, Ap
C 37	42.4	3.0	1512	US-09-252-991A-1551	Sequence 1551, Ap
C 38	42.4	3.0	1677	US-09-252-991A-1615	Sequence 1615, Ap
C 39	41.8	3.0	2025	US-09-252-991A-10293	Sequence 10293, A
40	41.8	3.0	2094	US-09-252-991A-10564	Sequence 10564, A
41	41.6	3.0	1425	US-09-540-236-972	Sequence 972, App
42	41.6	3.0	10528	US-09-902-540-945	Sequence 945, App
43	41.4	3.0	1164	US-09-902-540-5610	Sequence 5610, Ap
44	41.4	3.0	1206	US-09-902-540-4100	Sequence 4100, Ap
C 45	41.4	3.0	11812	US-09-902-540-1041	Sequence 1041, Ap

ALIGNMENTS

RESULT 1
US-09-252-991A-13529/c
; Sequence 13529, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13529
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-13529

Query Match	6.1%	Score 85.4;	DB 4;	Length 747;
Best Local Similarity	51.3%	Pred. No. 1.3e-12;		
Matches 250;	Conservative	0;	Mismatches 231;	Indels 6;
				Gaps 2;
Qy	651	GGTGAACCCGCGCGCGCAAAATGCGCTGCACGCTGAAGCCGACATTTTCGTTATTG	710	
Db	559	GGTGTCTTCCACCGCGCGCAATGTGCGGCATCTCCAGCGCGGCGATCGACGGTTCTG	500	
Qy	711	CGTGGCGAAGCGCTGGAACGACGAGCGCTGTTCAAGGAGATAGCGCTGCGTATCAC	770	
Db	499	CGCGCGGCGCGCTGGGGCGCGCTGGCGGTGACACAGGCGCAAGGCTTACATCGGCAC	440	
Qy	771	CGATGAAGAACTTGAAGACACCGCGGAAAAAGTTTGGGGGTGACCAACAAATGGC	830	
Db	439	CAACCGAGCGATGTGGCGGACCAACCGGAAAAAGTCTCTGGGACATACCGCGCTTCGT	380	
Qy	831	GGAGAAATACCCCAACATCTATCTGGGGGTGACCAAGCGCTGATTCGGCGGATCTG	890	
Db	379	CGAGCGCTTACCGCAACCGCGCGCGCGCGCTGTGTATGCG---GCTCTGACCGCGACCG	323	
Qy	891	GCTGAACCGCGCAATTAACAAGAACCGCAAGAGCCATGGAATGTGCGCGCAAAAACA	950	
Db	322	CTTCATTCGAACAGAAAGCGGAGAACCGCTGGGACACGCGCAACTGATCAGCGCGCGA	263	
Qy	951	ATACGTCGCGTCTGACGTGGAATGCTGGCGGCTACATGAAGGCACTTTCGAATPCCA	1010	
Db	262	CTATGTGACGCTCGCGCTGCGCGGATTCACGCGCGCTTCTTCTGGCGGCTACGAGACGG	203	

QY 1011 AAAAGACGATTAACGGCGCTACCGACTTCAACACCTTCTTTCGACAGGCGCCAGCTA 1070
DB 202 CCTCGGCAACGCTTGGAGGAGCCACATCCGCTCGCTTCTACGCGGCGCAAGTCAA 143
QY 1071 TCCGCTCT--ACAGAGTGCAGTCTGTATCTGATCCAGCTGAGCGCTGGGGCATGAT 1127
DB 142 CCGGCTTGGCTCTCGGATGGCATGTGGTTTCATGACCCAGTTCCGCGCTGGGGCTGTCT 83
QY 1128 CAATGAA 1134
DB 82 GCGCGAA 76

RESULT 2

US-09-252-991A-13930
; Sequence 13930, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13930
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13930

Query Match 6.1%; Score 85.4; DB 4; Length 750;
Best Local Similarity 51.3%; Pred. No. 1.3e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

QY 651 GGTGACCCCGCGCGCAATGCGTCAACGCTGGAAGCGGACCATTTTGGTTATTG 710
DB 54 GGTGTTTCCACCGCGCGCAATGATGATCGCATCTCCAGCGCGGCGCATGACGGGTTCTG 113
QY 711 GGTGCGGAGCGGTGGAACAGAGCGCGTGTTCAGAGGCGATAGGGTCCGGTATCAC 770
DB 114 GCG 173
QY 771 CGATGAAGACTCTGGAAGGACACGCGCGGAAAAAGTCTTCGCGGTGACCAAAATGAGC 830
DB 174 CAGCCAGGAGATCTGGCGCGGACCAAGGAAAAAGTCTTCGCGGTGACCAAGCGCGCTTCT 233
QY 831 GGAAGAAATACCCCAACCTATCTGCGGTGACCAAGCGCGGTGATTCGCGCGCGATCTG 890
DB 234 CAGCGCTACCGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 290
QY 891 GGTGAGCGCGGACAAATGAAGAACCGGCAAGAGCATGCAATCTGCGCGCAAAAAA 950
DB 291 CTTCATGGAACAGAAAGCGGAGAACCGCTGCGCAACCGGCAATCATGAGCGCGCGCA 350
QY 951 ATACGCTGCTGCTGAGCGTGAAGTGTGCGCGGTACATGAACGGCATTTTGAATACGA 1010
DB 351 CTATGTCAGCGCTCGCGCTCGCGCGCATCCAGCGCGCTTCTTCGCGCGGTACAGAGCG 410
QY 1011 AAAAGACGATTAACGGCGCTACCGGACTTCAACACCTTCTTTCGCGACGCGCGCACTA 1070
DB 411 CTTGCGCAACGCTTGGAGGAGCCACATCCGCTCGCTTCTACGCGGCGGAGATCAA 470
QY 1071 TCCGCTCT--ACAGAGTGCAGTCTGTATCTGATCCAGCTGAGCGCGTGGGGCATGAT 1127
DB 471 CCGGCTTGGCTCTCGGATGGCATGTGGTTTCATGACCCAGTTCCGCGCTGGGGCTGTCT 530
QY 1128 CAATGAA 1134

DB 531 GCGCGAA 537

RESULT 3

US-09-252-991A-13822
; Sequence 13822, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13822
; LENGTH: 831
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13822

Query Match 6.1%; Score 85.4; DB 4; Length 831;
Best Local Similarity 51.3%; Pred. No. 1.3e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

QY 651 GGTGACCCCGCGCGCAATGCGTCAACGCTGGAAGCGGACCATTTTGGTTATTG 710
DB 100 GGTGTTTCCACCGCGCGCAATGATGATCGCATCTCCAGCGCGGCGCATGACGGGTTCTG 159
QY 711 GGTGCGGAGCGGTGGAACAGAGCGCGTGTTCAGAGGCGATAGGGTCCGGTATCAC 770
DB 160 CCG 219
QY 771 CGATGAAGACTCTGGAAGGACACGCGCGGAAAAAGTCTTCGCGGTGACCAAAATGAGC 830
DB 220 CAGCCAGGAGATCTGGCGCGGACCAAGGAAAAAGTCTTCGCGGTGACCAAGCGCGCTTCT 279
QY 831 GGAAGAAATACCCCAACCTATCTGCGGTGACCAAGCGCGGTGATTCGCGCGCGATCTG 890
DB 280 CAGCGCTACCGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 336
QY 891 GGTGAGCGCGGACAAATGAAGAACCGGCAAGAGCATGCAATGCTGCGCGCAAAAAA 950
DB 337 CTTCATGGAACAGAAAGCGGAGAACCGCTGCGCAACCGGCAATCATGAGCGCGCGCA 396
QY 951 ATACGCTGCTGCTGAGCGTGAAGTGTGCGCGGTGACATGAACGGCATTTTGAATACGA 1010
DB 397 CTATGTCAGCGCTCGCGCTCGCGCGCATCCAGCGCGCTTCTTCGCGCGGTACAGAGCG 456
QY 1011 AAAAGACGATTAACGGCGCTACCGGACTTCAACACCTTCTTTCGCGACGCGCGCACTA 1070
DB 457 CTTGCGCAACGCTTGGAGGAGCCACATCCGCTCGCTTCTACGCGGCGGAGATCAA 516
QY 1071 TCCGCTCT--ACAGAGTGCAGTCTGTATCTGATCCAGCTGAGCGCGTGGGGCATGAT 1127
DB 517 CCGGCTTGGCTCTCGGATGGCATGTGGTTTCATGACCCAGTTCCGCGCTGGGGCTGTCT 576
QY 1128 CAATGAA 1134
DB 577 GCGCGAA 583

RESULT 4

US-09-489-039A-3192
; Sequence 3192, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KIEBSIELLA

;; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 2709.2004001
;; CURRENT APPLICATION NUMBER: US/09/489,039A
;; CURRENT FILING DATE: 2000-01-27
;; PRIOR APPLICATION NUMBER: US 60/117,747
;; PRIOR FILING DATE: 1999-01-29
;; NUMBER OF SEQ ID NOS: 14342
;; SEQ ID NO 3192
;; LENGTH: 1371
;; TYPE: DNA
;; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3192

Query Match 5.9%; Score 82.4; DB 4; Length 1371;
Best Local Similarity 56.2%; Pred. No. 9,76-12;
Matches 155; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 652 GTGACCCCGCCGCGCAAAATGCGTCAACGCTGGAACCGCGCAACATTTCGTTATTCG 711
DB 694 GTGGTCCGCGCGCGCGATGCTGATGAACATGCGCATTCGCAATATGAGCGCTTTTGC 753
QY 712 GTGCGCAGCGCGTGAACCAAGCGCGTGTTCAGAGGCGATGAGCGTGCCTGATCAC 771
DB 754 GTGCGCGAGCGCGTGAACCGCGCGCTATTATGACACCGTATCGCTTACCGCGGCACT 813
QY 772 GATGAAGACTCTGGAAGACACGCGCGAAAGTCTTTCGCGCTGACCAACATATGGCG 831
DB 814 TCCAGAGATATCGCGCGCGACATCCGAAAGTGTGTGGCAACCGTCCGACTGGGTG 873
QY 832 GAGAAATACCCCAACACTCTATCTGCGCGTGAACAGCGCGCTGATTCGGCGCGATCTCG 891
DB 874 GAACGCAACCCGCAACCGCGCGCGCTGTGTGGCGCGCTGATGAGCGCGCAAGCGCTGG 933
QY 892 CTGAGCGCGCAATATACAGAACCGCAAGAGAGCC 927
DB 934 ATCGCGCTTCGCGCGGAGAACCGCGGAGAGCGCC 969

RESULT 5
US-09-252-991A-13821
;; Sequence 13821, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OR INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 13821
;; LENGTH: 594
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13821

Query Match 5.0%; Score 69; DB 4; Length 594;
Best Local Similarity 52.2%; Pred. No. 2,46-08;
Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 84 GCCCGATGTGCGGCGCACTGCAAGCTGGAAGAAAGATTGAAATTCGCTTCATCAA 143
DB 204 GGCTTGCGCGCGGCTTCCAGCAGCAACGAAAGAGCCCTGACATCGTTTCATGCGC 263
QY 144 GCTCACCGACATGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203
DB 264 GCTGACCGATTCGCGCTCGCTGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 323
QY 204 CTGTTCGTGCAACTGGAAGCGCAACATGGAAGGTGTGTATGATGAGGTGCTGAA 263

DB 324 CTTACGCTAACTTCAGGCGCGCAACCGTCTTGCCCACTTGGCGCAAGCTGCTCAG 383
QY 264 TGGCAACTGGAAGGCTCGACATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 323
DB 384 CGGCGACTGATGTCGCGCGAGTCCGTATACGGCTGTGTACGGCTTCAGCGCTCGGCT 443
QY 324 CGGCAACCAAGCGCGATATCGAGGTGCGCTTCAGCANTGAGGCTTCAACGCGCAACG 376
DB 444 CGGCGGAGCGCGCGCAAGCATGAGTGGCGGTGCTCATGGAAGTGGCAAAACG 496

RESULT 6
US-09-252-991A-13627/c
;; Sequence 13627, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OR INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 13627
;; LENGTH: 3720
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13627

Query Match 5.0%; Score 69; DB 4; Length 3720;
Best Local Similarity 52.2%; Pred. No. 4,76-08;
Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 84 GCCCGATGTGCGGCGCACTGCAAGCTGGAAGAAAGATTGAAATTCGCTTCATCAA 143
DB 676 GGCTTGCGCGCGGCTTCCAGCAGCAACGAAAGAGCCCTGACATCGTTTCATGCGC 617
QY 144 GCTCACCGACATGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203
DB 616 GCTGACCGATTCGCGCTCGCTGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 557
QY 204 CTGTTCGTGCAACTGGAAGCGCAACGCGCAACTGGAAGGTGTGTATGAGGTGCTGAA 263
DB 556 CTTACGCTAACTTCAGGCGCGCAACCGTCTTGCGCGCAACCTTGGCGCAAGCTGCTCAG 497
QY 264 TGGCAACTGGAAGGCTCGACATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 323
DB 496 CGGCGACTGATGTCGCGCGAGTCCGTATACGGCTGTGTACGGCTTCAGCGCTCGGCT 437
QY 324 CGGCAACCAAGCGCGATATCGAGGTGCGCTTCAGCANTGAGGCTTCAACGCGCAACG 376
DB 436 CGGCGGAGCGCGCGCAAGCATGAGTGGCGGTGCTCATGGAAGTGGCAAAACG 384

RESULT 7
US-09-252-991A-9136
;; Sequence 9136, Application US/09252991A
;; Patent No. 6551795
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OR INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190

NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 777
LENGTH: 6655
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-777

Query Match 3.5%; Score 48.2; DB 4; Length 6655;
Best Local Similarity 44.3%; Pred. No. 0.017;
Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

```
QY 195 GACGAGGCGCTTGTGCACTGGAAGCGGCGCACTGGAAGGTGTGATGATG 254
DB 4813 GGTCAAGTGTGAGTATGTAACCGGCGCGCTTCATTAAGACCGCATGGGCTTACAT 4754
QY 255 GGTGTGATGAGCAACTGACCGCTTCGCAATGCTGGCGCGCGCTTGAAGCGGCGAG 314
DB 4753 CCTCGAAGAGCGGAGCGGAGGGAAGCTCAAGCCCGGCGGACCATGCTCGAAGAAC 4694
QY 315 CGTTGCTTGGCAACCAAGCGCATATGAGTGCCTTACGATGGGCTTCAACGGCAA 374
DB 4693 GTCCGCAACACCGGCGCATGCGCGCTGGCGCGCGCGCTCAAGGCTTAAAGTGCAT 4634
QY 375 CGGATTAACGCTGCAATGAATCTGGCATGAGTGAAGCCGAACATACCGCTGGAAG 434
DB 4633 CTTCACCATGCGGACCAAGATGTCCTTGAAGAGATCAACCGCTCAAGCGCTGGGCGC 4574
QY 435 CGGTAAACCGGTGATCGATCAAGGAGATTAATCTAAACCGGTGCTGCAAAAATCA 494
DB 4573 GCAAGTGTGTGATGACCGCGCAAGACGTGCGCGCGGAGAGACCGCGGACGTACGAGAC 4514
QY 495 AGCCGAAGGAGAGCGGCTTCAATATGCGATGACCTTCCGCGCGGATCTCAACATCAA 554
DB 4513 GGCACAGCGCTGACCGGAGAGACGCGGCGCGCTTATGCTCAACGATACCAACCC 4454
QY 555 ACTGCGTATGCTGCGCGGCTGCGGATCAATCTGCTGCTATTCGCGCGGAGGA 614
DB 4453 CGACCAATGAGCGCGCATCAACACACCGGCTGAGATTAACGAAGACCGAGGG 4394
QY 615 CATTCGCGGCAATCGGCGGAGAC 639
DB 4393 CAAGTTCGACTACTTCTGCGGCGC 4369
```

RESULT 11

US-09-902-540-5834
Sequence 5834, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 5834
LENGTH: 750
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-5834

Query Match 3.4%; Score 48; DB 4; Length 750;
Best Local Similarity 47.9%; Pred. No. 0.0086;
Matches 138; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```
QY 144 GCTCACCGAATGAGCGCGCTGCGGCTGCGCGCGGAAAAAGCTTTTCAGACGAGGG 203
DB 204 GCGGCGGAGATTGCGCGAGGTGACGTGTGTCAACACGCGGCGCTGCGGCGGTT 263
```

```
QY 204 CTTGTTGTCGCACTGGAAGCGGCAACCTGGAAGGTGTGATGATGAGGTGCGTGA 263
DB 264 GACCTGGCGCAATCGACCGCGCTGAGAGCTGGACACGATGTGACCAACGAGAA 323
QY 264 TGGCAACTGAGCGGCTCGACATGCTGGGCGCGCGCGCTTACGCGCACGCTTGCTT 323
DB 324 GGGCTTCCTTACTGACGACGACCGGCTGCTGCCCGGATGTGTGGCGGCAACCGGGCCA 383
QY 324 CGGCAACGAGCGGATATCGAGGTGCGCTTACGATGGCTTCAACGCGCAACCGATTAC 383
DB 384 CGTGTCAACATGAGGCTCGGTGGCGGAGTGGCCCTACCGGCGGCAACGTTACGG 443
QY 384 GGTGTCAATGAATCTGGCATGATGAAGCCGAACATACCGCTGA 431
DB 444 CGCACCAAGCGCTTGTGTGACCAAGTTCAGCTTCAACCTCGCGCGA 491
```

RESULT 12

US-09-902-540-1273/c
Sequence 1273, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1273
LENGTH: 72704
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(72704)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1273

Query Match 3.4%; Score 48; DB 4; Length 72704;
Best Local Similarity 47.9%; Pred. No. 0.046;
Matches 138; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```
QY 144 GCTCACCGAATGAGCGCGCTGCGGCTGCGCGCGGAAAAAGCTTTTCAGACGAGGG 203
DB 70204 GCGGCGGAGATTGCGCGAGGTGACGTGTGTCAACACGCGGCGCTGCGGCGGTT 70145
QY 204 CTTGTTGTCGCACTGGAAGCGGCAACCTGGAAGGTGTGATGATGAGGTGCGTGA 263
DB 70144 GACCTGGCGCAATCGACCGCGCTGAGAGCTGGACACGATGTGACCAACGAGAA 70085
QY 264 TGGCAACTGAGCGGCTCGACATGCTGGGCGCGCGCGCTTACGCGCACGCTTGCTT 323
DB 70084 GGGCTTCCTTACTGACGACGACCGGCTGCTGCCCGGATGTGTGGCGGCAACCGGGCCA 70025
QY 324 CGGCAACGAGCGGATATCGAGGTGCGCTTACGATGGCTTCAACGCGCAACCGATTAC 383
DB 70024 CGTGTCAACATGAGGCTCGGTGGCGGAGTGGCCCTACCGGCGGCAACGTTACGG 69965
QY 384 GGTGTCAATGAATCTGGCATGATGAAGCCGAACATACCGCTGA 431
DB 69964 CGCACCAAGCGCTTGTGTGACCAAGTTCAGCTTCAACCTCGCGCGA 69917
```

RESULT 13

US-08-592-874-1/c
Sequence 1, Application US/08592874
Patent No. 5854034

GENERAL INFORMATION:
APPLICANT: POLLOCK, THOMAS J.
APPLICANT: YAMAZAKI, MOTOHIDE
APPLICANT: THORNE, LINDA
APPLICANT: MIKOLAJCZAK, MARCIA
APPLICANT: ARMENTROUT, RICHARD W.
TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: JULES E. GOLDBERG
STREET: 261 MADISON AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,874
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/377,440
FILING DATE: 24-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: GOLDBERG, JULES E.
REGISTRATION NUMBER: 24,408
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-986-4090
TELEFAX: 212-818-9479
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 28804 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FRAGMENT TYPE: N-terminal
US-08-592-874-1

Query Match 3.4%; Score 47.8; DB 2; Length 28804;
Best Local Similarity 47.0%; Pred. No. 0.037;
Matches 148; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 64 GTTTGGGGTTTGACATGTCGCGGCGAGTTGGCAAGCTGGAAAAAGAAAGT 123
DB 18736 GTGGCGGGCTGGCATCACCGCTGGCTGGCAATCGCGCGTGAAGTGTGGCG 18677
QY 124 TTGAATTCGGCTTCAAGCTCACCGCATGCGCCGCTGGCGGTGGCCGCAAAA 183
DB 18676 CGAGCGAGGGCGCATGCGCCGCGATCGGCGAGACCAAGATGTGAGTCCCGCAAGC 18617
QY 184 GCGTTTTCGAGAGCAGAGGCGCTGTTCTGTGCACTGGAAGCGCAGCCCACTGGAAGTGTG 243
DB 18616 GGTATGTCGCGGCGCATTTCTGTGTGGCGCAGAGGCGTTCAGAAAGGCCAGGTGTCTG 18557
QY 244 GTGATGATAGAGGTGCTGATGAGTGGCACTGGAACGCGTCCGCAATGTGCGCGCGCGCGCG 303
DB 18556 ATCAAGCTGTGATCCCACTGTGCGAGCCGGAAGCGCGCGCGTGGGCTGTCTC 18497
QY 304 TTAGCGGCGCAGGCTTTCGCGCACCAAGCGCGATATGAGGTGCGTTTCAGCATGGGC 363
DB 18496 AGCGCCAGCTGATCCCGCGCGCACAGCGCATATGACGCGCTGAGACGGAAGGCGC 18437
QY 364 TTCACGCGCAACGCG 378
DB 18436 TTCGCTTCGCGCG 18422

RESULT 14
US-09-096-942-2/c
Sequence 2, Application US/09096942
Patent No. 6027925
GENERAL INFORMATION:
APPLICANT: Pollock, Thomas J
APPLICANT: Mikolajczak, Marcia
APPLICANT: Yamazaki, Motohide
APPLICANT: Thorne, Linda
APPLICANT: Armentrout, Richard W
TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
FILE REFERENCE: seq list for appl filed from pro. appl
CURRENT APPLICATION NUMBER: US/09/096,942
EARLIER FILING DATE: 1998-06-12
EARLIER APPLICATION NUMBER: 60/049,428
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 28804
TYPE: DNA
ORGANISM: Sphingomonas sp. S88
US-09-096-942-2

Query Match 3.4%; Score 47.8; DB 3; Length 28804;
Best Local Similarity 47.0%; Pred. No. 0.037;
Matches 148; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 64 GTTTGGGGTTTGACATGTCGCGGCGAGTTGGCAAGCTGGAAAAAGAAAGT 123
DB 18736 GTGGCGGGCTGGCATCACCGCTGGCTGGCAATGCGCGCGTGAAGTGTGGCG 18677
QY 124 TTGAATTCGGCTTCAAGCTCACCGCATGCGCCGCTGGCGGTGGCCGCAAAA 183
DB 18676 CGAGCGAGGGCGCATGCGCCGCGATCGGCGAGACCAAGATGTGAGTCCCGCAAGC 18617
QY 184 GCGTTTTCGAGAGCAGAGGCGCTGTTCTGTGCACTGGAAGCGCAGCCCACTGGAAGTGTG 243
DB 18616 GGTATGTCGCGGCGCATTTCTGTGTGGCGCAGAGGCGTTCAGAAAGGCCAGGTGTCTG 18557
QY 244 GTGATGATAGAGGTGCTGATGAGTGGCACTGGAACGCGTCCGCAATGTGCGCGCGCGCGCG 303
DB 18556 ATCAAGCTGTGATCCCACTGTGCGAGCCGGAAGCGCGCGCGCTGGGCTGTCTC 18497
QY 304 TTAGCGGCGCAGGCTTTCGCGCACCAAGCGCGATATGAGGTGCGTTTCAGCATGGGC 363
DB 18496 AGCGCCAGCTGATCCCGCGCGCACAGCGCATATGACGCGCTGAGACGGAAGGCGC 18437
QY 364 TTCACGCGCAACGCG 378
DB 18436 TTCGCTTCGCGCG 18422

RESULT 15
US-09-096-967-2/c
Sequence 2, Application US/09096867
Patent No. 6030817
GENERAL INFORMATION:
APPLICANT: Pollock, Thomas J
APPLICANT: Mikolajczak, Marcia
APPLICANT: Yamazaki, Motohide
APPLICANT: Thorne, Linda
APPLICANT: Armentrout, Richard W
TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
FILE REFERENCE: seq list for appl filed from pro. appl
CURRENT APPLICATION NUMBER: US/09/096,867
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 60/049,428
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Query Match 3.4%; Score 47.8; DB 3; Length 28804;
Best Local Similarity 47.0%; Pred. No. 0.037;
Matches 148; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```
QY 64 GTTGGGGTTTGAAGATTGGCCGATGTGGGGCAGTTGGCAAGCTGAAAAAGAAAT 123
    |||||
DB 18736 GTGGCGGGGCTGGCGATCACCACCGCTGGCTGGCAATCGCCGCTGGAATGTGTGGCG 18677
    |||||
QY 124 TTGAAATTGGCTTTCATCAAGCTCACCAGATGGCGCGCTGGCGGTGGCCGCAAAAA 183
    |||||
DB 18676 CCGACGAGGGCGCGCATCGCCCGATGGCGAGACCAAGATCGTGCAGTCCCCGAAAGC 18617
    |||||
QY 184 GCGTTCTTGAAGACGAGGGCTGTTGTGCAACTGGAAGCGCAAGCCCAACTGGAAGGTG 243
    |||||
DB 18616 GGTATGTCGCGCGCATTTGTGTGGGCGAGGGCGAAGGTTTCAGAAAGGCGCAGGTGCTG 18557
    |||||
QY 244 GTGATGATAGGGTTCGTAATGGGAACTGGACGGCTCGCACATGTGGCGCGGCGCGCG 303
    |||||
DB 18556 ATCAGCTCGATCCACCATGTGGCGACGCCGAAGCCGCGCAGGCGCGCTGTGCTTC 18497
    |||||
QY 304 TTAGCGGCCAGCGTTGGCTTCGGCACCAGCCGATATCGAGGTGCCGTTCAAGCATGGGC 363
    |||||
DB 18496 AGGCGCCAGCTCGATGCGCGCGGGAACCAAGCGATCATCGACGCGCTGGACGGAAGGCG 18437
    |||||
QY 364 TTCACGGGCAAGCGG 378
    |||||
DB 18436 TTCGCTTCGCGCGG 18422
    |||||
```

Search completed: May 20, 2005, 23:22:36
Job time : 261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 20, 2005, 19:59:08 ; Search time 889 Seconds
(without alignments)
9602.015 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaacaccatcatgatac.....ccgcgcgaagtcgtgat 1392

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5695437 seqs, 3066160638 residues

Total number of hits satisfying chosen parameters: 11390874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392	100.0	1392	US-10-689-200-1	Sequence 1, Appl1
2	143	10.3	551	US-10-689-200-3	Sequence 3, Appl1
3	71.8	5.2	1209	US-10-282-122A-33791	Sequence 33791, A
4	62.2	4.5	1173	US-10-282-122A-11998	Sequence 11998, A
5	59.2	4.3	576	US-10-282-122A-11492	Sequence 11492, A
6	58	4.2	1200	US-10-282-122A-30186	Sequence 30186, A
7	58	4.2	1200	US-10-389-647-149	Sequence 149, App
8	57	4.1	1116	US-10-282-122A-33179	Sequence 32179, A
9	49.4	3.5	1512	US-10-369-493-42267	Sequence 42267, A
10	48.4	3.5	773	US-10-282-122A-8343	Sequence 8343, Ap
11	48	3.4	744	US-10-369-493-43243	Sequence 43243, A

12	47.8	3.4	1017	US-10-437-963-85118	Sequence 85118, A
13	47.8	3.4	7656	US-10-282-122A-25488	Sequence 25488, A
14	46.4	3.3	3048	US-10-156-761-3146	Sequence 3146, Ap
15	46.4	3.3	9025608	US-10-156-761-1	Sequence 1, Appl1
16	46.2	3.3	1440	US-10-437-963-82648	Sequence 82648, A
17	45.6	3.3	1173	US-10-282-122A-13773	Sequence 13773, A
18	45.4	3.3	1396	US-10-767-701-13423	Sequence 13423, A
19	44.4	3.2	930	US-10-437-963-93713	Sequence 93713, A
20	44.4	3.2	1365	US-10-437-963-56913	Sequence 56913, A
21	44	3.2	786	US-10-437-963-13840	Sequence 13840, A
22	43.4	3.1	1188	US-10-282-122A-13623	Sequence 13623, A
23	43.4	3.1	5577	US-10-369-493-43212	Sequence 43212, A
24	43.2	3.1	1155	US-10-437-963-67845	Sequence 67845, A
25	43.2	3.1	2110	US-10-437-963-10628	Sequence 10628, A
26	43	3.1	1305	US-10-282-122A-11478	Sequence 11478, A
27	43	3.1	1818	US-10-369-493-27231	Sequence 27231, A
28	43	3.1	2021	US-10-437-963-49750	Sequence 49750, A
29	42.8	3.1	956	US-10-425-114-24149	Sequence 24149, A
30	42.8	3.1	984	US-10-425-115-118414	Sequence 118414, A
31	42.8	3.1	993	US-10-437-963-93714	Sequence 93714, A
32	42.6	3.1	2213	US-10-772-636-27	Sequence 27, Appl1
33	42.6	3.1	2931	US-10-282-122A-15175	Sequence 15175, A
34	42.4	3.0	1132	US-10-156-761-4985	Sequence 4985, Ap
35	42.4	3.0	1305	US-10-282-122A-12878	Sequence 12878, A
36	42.4	3.0	1470	US-09-815-242-7855	Sequence 7855, Ap
37	42.4	3.0	9025608	US-10-156-761-1	Sequence 1, Appl1
38	42.2	3.0	375	US-10-425-114-4691	Sequence 4691, Ap
39	42.2	3.0	501	US-10-425-114-319	Sequence 319, App
40	42.2	3.0	726	US-10-425-114-58	Sequence 58, Appl
41	42.2	3.0	1525	US-10-369-493-27882	Sequence 27882, A
42	42.2	3.0	1866	US-10-425-115-74324	Sequence 74324, A
43	42	3.0	536	US-10-338-110-119	Sequence 119, App
44	41.6	3.0	900	US-10-369-493-41628	Sequence 41628, A
45	41.6	3.0	1271	US-10-437-963-38282	Sequence 38282, A

ALIGNMENTS

RESULT 1
US-10-689-200-1
Application US/10689200
Publication No. US20040126848A1
GENERAL INFORMATION:
APPLICANT: DiCosimo, Deana J.
APPLICANT: Ni, Hao
APPLICANT: Ye, Rick
APPLICANT: Pictagallo, Stephen
APPLICANT: Wang, Tao
APPLICANT: Seid, John E.
TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN CL METABOLIZING BACTERIA
FILE REFERENCE: C11747 US NA
CURRENT APPLICATION NUMBER: US/10/689,200
PRIOR FILING DATE: 2003-10-20
PRIOR APPLICATION NUMBER: 60/419,872
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1392
TYPE: DNA
ORGANISM: *Mechyliomonas* sp. 16a
US-10-689-200-1

Query Match 100.0%; Score 1392; DB 18; Length 1392;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAACCATCTTGGATCGAGCTGCAAGAAATTTGTTATGACGTTATCGGCTGCCTA 60
DB 1 ATGAAACCATCTTGGATCGAGCTGCAAGAAATTTGTTATGACGTTATCGGCTGCCTA 60
QY 61 GCCTTGGGCTTTCAGATTGCGCCCGATGTCGGGGCACTGGCAAGTGAAGAAAGAA 120

Db 61 GCGGTTGGGGTTGACGATTCGCGCCGATGTCGGGCACTTGGCAAGCTGGAAAAAGAA 120
Qy 121 GATTGAAATTCGAGCTTCATCAAGCTCAACGACATGGCGCGCTGGCGATGGCGCGAA 180
Db 121 GATTGAAATTCGAGCTTCATCAAGCTCAACGACATGGCGCGCTGGCGATGGCGCGAA 180
Qy 181 AAGGCTTCCTTGAGGACGAGGGGCTGTTGTTGCAACTGGAAACGACGCGCACTGGAA 240
Db 181 AAGGCTTCCTTGAGGACGAGGGGCTGTTGTTGCAACTGGAAACGACGCGCACTGGAA 240
Qy 241 GTGGTATGATAGGAGTGTGTAATGGCGAATGGACGCTCGCAATGCTGGCGCGCG 300
Db 241 GTGGTATGATAGGAGTGTGTAATGGCGAATGGACGCTCGCAATGCTGGCGCGCG 300
Qy 301 CCGTTAGCGGCGACGCTTGGCTTCGCGACCAAGGCGGATTCGAGGTGCGGTTCAAGAT 360
Db 301 CCGTTAGCGGCGACGCTTGGCTTCGCGACCAAGGCGGATTCGAGGTGCGGTTCAAGAT 360
Qy 361 GCGTTAAGCGGCAACCGGATTCGAGGTTCGCAATGAATCTGGCATCAGATGAAGCGGAA 420
Db 361 GCGTTAAGCGGCAACCGGATTCGAGGTTCGCAATGAATCTGGCATCAGATGAAGCGGAA 420
Qy 421 ATACCGCTGGAAGCGGATTAACCGGTGCATCCGATCAAGGACAGATTATCTAAACCGGTC 480
Db 421 ATACCGCTGGAAGCGGATTAACCGGTGCATCCGATCAAGGACAGATTATCTAAACCGGTC 480
Qy 481 GTGGAATAATCAAGCGGCAAGCGGCAAGCGGTCATATGCGATGACCTTCCGCGCGGA 540
Db 481 GTGGAATAATCAAGCGGCAAGCGGCAAGCGGTCATATGCGATGACCTTCCGCGCGGA 540
Qy 541 TCTCAACAATCAAGCGGCTTCTGCTGCGGCGGCTGCGGATTCGATCTGCTGCTAT 600
Db 541 TCTCAACAATCAAGCGGCTTCTGCTGCGGCGGCTGCGGATTCGATCTGCTGCTAT 600
Qy 601 TCGCCCGCGGAGCACTTTCGCGCAAAATCGGCGGCAAGCGCTTGTGTCGCGGACCGG 660
Db 601 TCGCCCGCGGAGCACTTTCGCGCAAAATCGGCGGCAAGCGCTTGTGTCGCGGACCGG 660
Qy 661 CCGCGCGCAATGCGCTGCAAGCGGCAAGCGGCAAGCGGCAATTCGTTATGTCGCGGAG 720
Db 661 CCGCGCGCAATGCGCTGCAAGCGGCAAGCGGCAAGCGGCAATTCGTTATGTCGCGGAG 720
Qy 721 CCGTGAACGAGGCGGCTTCAAGGCGATGAGCGGCTGCGGATTCAGCCGATGAAGAA 780
Db 721 CCGTGAACGAGGCGGCTTCAAGGCGATGAGCGGCTGCGGATTCAGCCGATGAAGAA 780
Qy 781 CTCTGGAAGGACGCGCGGAAAAAGTCTTCGCGGTGACCAACCAATGCGCGGAGAAATAC 840
Db 781 CTCTGGAAGGACGCGCGGAAAAAGTCTTCGCGGTGACCAACCAATGCGCGGAGAAATAC 840
Qy 841 CCCAACAATCTATGCGGCTGCAAGGCGGCTGATTCGCGGCGCGATTCGCTGACGCGC 900
Db 841 CCCAACAATCTATGCGGCTGCAAGGCGGCTGATTCGCGGCGCGATTCGCTGACGCGC 900
Qy 901 GACAAATTAAGAAACGCGCAAGGAGCAATCGAATCTGCGCGGAAAAAATAATCTGCGGT 960
Db 901 GACAAATTAAGAAACGCGCAAGGAGCAATCGAATCTGCGCGGAAAAAATAATCTGCGGT 960
Qy 961 GCTGAGTGAAGTGTGCGGCTGACATGAACGCGCACTTTCGAAATCAAGAAAAAGCGAT 1020
Db 961 GCTGAGTGAAGTGTGCGGCTGACATGAACGCGCACTTTCGAAATCAAGAAAAAGCGAT 1020
Qy 1021 AAAACGCGCTACCGGACTTCACACCTTCTTTCGCGCAAGCGCGCAGCTATCGGTCTAC 1080
Db 1021 AAAACGCGCTACCGGACTTCACACCTTCTTTCGCGCAAGCGCGCAGCTATCGGTCTAC 1080
Qy 1081 AAGAGTGAAGTGTGATTCGACCCGAGCTGAGGCGCTGGGCGCATGATGAATTCGAA 1140
Db 1081 AAGAGTGAAGTGTGATTCGACCCGAGCTGAGGCGCTGGGCGCATGATGAATTCGAA 1140
Qy 1141 CCGGACAACTGGATTCGATACCGCGCAAGCGTTCACCGCGCGGACATCTATCGCGC 1200
Db 1141 CCGGACAACTGGATTCGATACCGCGCAAGCGTTCACCGCGCGGACATCTATCGCGC 1200

Db 1141 CCGGACAACTGGATTCGATACCGCGCAAGCGTTCACCGCGCGGACATCTATCGCGC 1200
Qy 1201 GCGGCCAAGAACTGTGTCGAGAGGCAAGGCCAAGGCGGAAAGCTTCCCTGCGATACC 1260
Db 1201 GCGGCCAAGAACTGTGTCGAGAGGCAAGGCCAAGGCGGAAAGCTTCCCTGCGATACC 1260
Qy 1261 TCGATCAAGCGGCTGAGAAATTTCTTCATGCAAAAGTCCGTTGATGCAACGAGGCC 1320
Db 1261 TCGATCAAGCGGCTGAGAAATTTCTTCATGCAAAAGTCCGTTGATGCAACGAGGCC 1320
Qy 1321 AACGATTACCTGCGCAAGTTTGGATAGGTCTGAAAGGCAAGCAACCGTAGCGCGGCG 1380
Db 1321 AACGATTACCTGCGCAAGTTTGGATAGGTCTGAAAGGCAAGCAACCGTAGCGCGGCG 1380
Qy 1381 AAGTGTGAT 1392
Db 1381 AAGTGTGAT 1392

RESULT 2
US-10-689-200-3
; Sequence 3, Application US/10689200
; Publication No. US20040126848A1
; GENERAL INFORMATION:
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Ni, Hao
; APPLICANT: Ye, Rick
; APPLICANT: Picatagallo, Stephen
; APPLICANT: Wang, Tao
; APPLICANT: Seip, John E.
; TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN CL METABOLIZING BACTERIA
; FILE REFERENCE: CL1747 US NA
; CURRENT APPLICATION NUMBER: US/10/689, 200
; PRIOR FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 60/419, 872
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Methylobionas sp. 16a
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (409)..(411)
; OTHER INFORMATION: ATG start site of mra gene
US-10-689-200-3

Query Match 10.3%; Score 143; DB 18; Length 551;
Best Local Similarity 100.0%; Pred. No. 7.1e-34;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAACCATATTAGATCGAGCTCGAAGAAATTTGATGACGTTATCGGCTTCGCTA 60
Db 409 ATGAAAAACCATATTAGATCGAGCTCGAAGAAATTTGATGACGTTATCGGCTTCGCTA 468
Qy 61 GCGGTTGGGTTTGAAGATTGCGCGCGGATGTCGCGGCGAGTTGGCGAAGTGAAGAAAGAA 120
Db 469 GCGGTTGGGTTTGAAGATTGCGCGCGGATGTCGCGGCGAGTTGGCGAAGTGAAGAAAGAA 528
Qy 121 GATTGAAATTCGCTTCATCA 143
Db 529 GATTGAAATTCGCTTCATCA 551

RESULT 3
US-10-282-122A-33791
; Sequence 33791, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

Query Match	5.2%	Score 71.8	DB 17	Length 1209
Best Local Similarity	53.8%	Pred. No. 2e-11		
Matches 148	Conservative 0	Mismatches 127	Indels 0	Gaps 0
QY	102	TGCGACGTGAAAAAGAAATTGTAATTCGGCTTCATCAAGTCACCGACATGGCC	161	
Db	45	TGACGCCCGGAAAAAGAACAGCTTGGAGCTGGCTTCATGGCACTGACCGACTGGCCCC	104	
QY	162	GCTGGCGGTGGCCCGCGGAAAAAGCTTCTTCGAGGACGAGGGCTGTTCGTGCACTGGA	221	
Db	105	GCTGGGTGCTGCTGTCACCGAGGGTTTGCCACGCTTAACGGGCTGAGGCTGAACCTGAA	164	
QY	222	AGCGGAGCCCAACTGGAAGGTGTGATGATGTGGTTCGTGAATGGCGCAACTGGAAGGCTCT	281	
Db	165	ACGCCAGACCTCTTGGGCCGGGCTCGTGAACAGGCTGTGACGCGGGCAGCTTCAAGCGGCG	224	
QY	282	GCAATGTGGGCGCGGCGCGCTTATGCGGCCAGCGTTGGCTTCGGCAACMAAGCCGATAT	341	
Db	225	ACATAGCCTGTATGTGTTGATTATATGACCGTGGAGCTGGGCAATCAACGGCGGCCCGCGAC	284	
QY	342	CGAGGTGCGTTACGACATGGGCTTCAACGGCAAG	376	
Db	285	GCAATGGCCATCTCATGTGGGCTTCAACCGAAGC	319	

	Query Match	Similarity	5.4%;	Score 62.2;	DB 17;	Length 1173;
	Best Local	Similarity	48.4%;	Pred. No. 2e-08;		
	Matches 128;	Conservative	0;	Mismatches 88;	Indels 3;	Gaps 1;
QY	131	TCGGCTTCATCAGAGCTACCGACATGGAGCGCGCTGGCGGTGGCCGCCGAAAGAGCTTCT	190			
Db	116	TCGGCTTCATCGCCATCACCGAGCCGGCCCGCTGTGCTGCCCAACACACGGCTATT	175			
QY	191	TCGAGGACGAGGGGCTGT---TCGTGCACATCGAAGGCGAGGCCAACATCGAAGGTGGTGA	247			
Db	176	TCGGGTCCGAAGGCTACACGCTGACGACGACCGAAGCTGTGCTGAGTGTGGCCGACGCTCG	235			
QY	248	TGAGTAGGGTTCGTAAATGCGAAGCTGACCGGCTTCGACATGCTTGGCGCCCGGCGCGTTAG	307			
Db	236	TCGAGGAGTTCTGTCTCGGCGCAAGTCAACGTCGTGCACCTGTCTGCCCGATGACGCTGT	295			
QY	308	CGGCGAGCGTTGGGTGGGCAACCAAGGCCATATCGAGG	346			
Db	296	GAGCGCGCTACGAGCAGCCGCGCCCGGCGAAGGTGGTGG	334			

```
RESULT 5
US-10-282-1122A-11492
: Sequence 11492, Application US/10282122A
: Publication No. US20040029129A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
```

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 11492
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-10-282-122A-11492

```

```

Query Match      4.3%; Score 59.2; DB 17; Length 576;
Best Local Similarity 57.9%; Pred. No. 1.3e-07;
Matches 155; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

```

```

QY 131 TCGGCTTCAATCAAGCTACCGACATGCGCGCGTGGCGCGCCGAAAGGCTTCT 190
    |||||
DB 116 TCGGCTATCTGCGGATCAACGACGCGCGCGTGTGTGCTGCTGCAACACGCTATT 175
    |||||
QY 191 TCGAGACGAGGCGCTGT--TGTGCAACTGAGAGCGCAACTGGAAGTGTGA 247
    |||||
DB 176 TCGCGCTCCAGAGGCTGACGCTGACCAACCCAAAGCTGCTGCGCACTGGCGCACTG 235
    |||||
QY 248 TGGATAGGGTCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 307
    |||||
DB 236 TCGAGGCTTCTCTGCTGCGGCAAGTCAAGTCTGCAAGTCTGCTGCTGCTGCTGCTGCT 295
    |||||
QY 308 CGGCGAGCGTTGGCTTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 343
    |||||
DB 296 GGGCAGCTACGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 331
    |||||

```

RESULT 6

```

US-10-282-122A-30186
; Sequence 30186, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari

```

```

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30186
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-30186

```

```

Query Match      4.2%; Score 58; DB 17; Length 1200;
Best Local Similarity 56.2%; Pred. No. 4.1e-07;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

```

```

QY 131 TCGGCTTCAATCAAGCTACCGACATGCGCGCGTGGCGCGCCGAAAGGCTTCT 190
    |||||
DB 137 TCGGCTACTGCGGATCAACGACGCGCGCTTGTGTGCTGCTGCAACGCGCTGT 196
    |||||
QY 191 TCGAGACGAGGCGCTGTGCTGCAACTGGAAGCGCAGGCC--AACTGGAAGTGTGA 247
    |||||
DB 197 TCGAGGCTCCAGGCGATCCAGCGCGGCAACGCGCGTGTGCTGCTGCTGCTGCTGCTGCT 256
    |||||
QY 248 TGGATAGGGTCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 307
    |||||
DB 257 TCGAGGCTTCTCTGCTGCGGCAAGTCAAGTCTGCAAGTCTGCTGCTGCTGCTGCTGCT 316
    |||||
QY 308 CGGCGAGCGTTGGCTTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 367
    |||||
DB 317 GGGCC-----CGTACGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 370
    |||||
QY 368 ACGGCAAGCGGATTAAGTGTTCATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 399
    |||||
DB 371 GCGGCTCGGGGCTCAAGTGTGCTCGCGAGATC 402
    |||||

```

RESULT 7

```

US-10-389-647-149
; Sequence 149, Application US/10389647
; Publication No. US20040033549A1
; GENERAL INFORMATION:
; APPLICANT: GREENBERG, E. Peter
; APPLICANT: SCHUSTER, Martin
; APPLICANT: LOSTROH, Candi

```

;; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
;; FILE REFERENCE: U12-038CP
;; CURRENT APPLICATION NUMBER: US/10/389,647
;; CURRENT FILING DATE: 2003-03-14
;; PRIOR APPLICATION NUMBER: 09/653730
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/153022
;; PRIOR FILING DATE: 1999-09-03
;; NUMBER OF SEQ ID NOS: 710
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 149
;; LENGTH: 1200
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-10-389-647-149

Query Match 4.2%; Score 58; DB 17; Length 1200;
Best Local Similarity 56.2%; Pred. No. 4.1e-07;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

Qy 131 TCGGCTTCATCAAGCTCACCGACATGCGCGCGCTGGCGCGCCGAAAGGCTTCT 190
Db 137 TCGGCTTCATCAAGCTCACCGACATGCGCGCGCTGGCGCGCCGAAAGGCTTCT 196
Qy 191 TCGAGGACGAGGCGCTGTTCTGTCGCACTGGAAGCGGAGCC--AACTGGAAGTGTGTA 247
Db 197 TCGAGGCGGAGGCGCTGTTCTGTCGCACTGGAAGCGGAGCC--AACTGGAAGTGTGTA 256
Qy 248 TGGATAGGCTGCGTAATGCGGAATGGAAGTGGAGCGGCTGCGACATGCTGGCGCGCCGCTTAG 307
Db 257 TCGAGGCGGCTGCGTAATGCGGAATGGAAGTGGAGCGGCTGCGACATGCTGGCGCGCCGCTTAG 316
Qy 308 CGGCGACGCTTGGCTTCGCGACCAAGCGCGATATCGAGTGGCGCTTCAGCATGGGCTTCA 367
Db 317 GGGCC-----CGCTACGGAGCAAGATGCGCGCGCAAGGTGGCTGGAACACAGCTCG 370
Qy 368 ACGGCAACGCGATTACGCTGTCTCCATGAAATC 399
Db 371 GCGGCTCGGGGCTCACGCTGTGCTCCGAGATC 402

RESULT 8
US-10-282-122A-32179

;; Sequence 32179, Application US/10282122A
;; Publication No. US20040029129A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Liangu
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Malone, Cheryl
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl
;; APPLICANT: Zyskind, Judith
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John
;; APPLICANT: Carr, Grant
;; APPLICANT: Yamamoto, Robert
;; APPLICANT: Forsyth, R.
;; APPLICANT: Xu, H.
;; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;; FILE REFERENCE: ELITRA.034A
;; CURRENT APPLICATION NUMBER: US/10/282,122A
;; CURRENT FILING DATE: 2003-02-20
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578

;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 32179
;; LENGTH: 1116
;; TYPE: DNA
;; ORGANISM: Pseudomonas putida
US-10-282-122A-32179

Query Match 4.1%; Score 57; DB 17; Length 1116;
Best Local Similarity 54.1%; Pred. No. 8.2e-07;
Matches 139; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

Qy 131 TCGGCTTCATCAAGCTCACCGACATGCGCGCGCTGGCGCGCCGAAAGGCTTCT 190
Db 56 TCGGCTTCATCAAGCTCACCGACATGCGCGCGCTGGCGCGCCGAAAGGCTTCT 115
Qy 191 TCGAGGACGAGGCGCT--GTTCTGTCGCACTGGAAGCGGAGCCAACTGGAAGTGTGTA 247
Db 116 TCGAGGCGGAGGCGCT--GTTCTGTCGCACTGGAAGCGGAGCCAACTGGAAGTGTGTA 175
Qy 248 TGGATAGGCTGCGTAATGCGGAATGGAAGTGGAGCGGCTGCGACATGCTGGCGCGCCGCTTAG 307
Db 176 TCGAGGCGGAGGCGCTGCGTAATGCGGAATGGAAGTGGAGCGGCTGCGACATGCTGGCGCGCCGCTTAG 235
Qy 308 CGGCGACGCTTGGCTTCGCGACCAAGCGCGATATCGAGTGGCGCTTCAGCATGGGCTTCA 367
Db 317 GGGCC-----CGCTACGGAGCAAGATGCGCGCGCAAGGTGGCTGGAACACAGCTCG 370
Qy 368 ACGGCAACGCGATTACG 384
Db 296 GCGGCTGACTGTGGCG 312

RESULT 9
US-10-369-493-42267

;; Sequence 42267, Application US/10369493
;; Publication No. US20030233675A1
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Chen, Xianning
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 42267
;; LENGTH: 1512
;; TYPE: DNA
;; ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-42267

Query Match 3.5%; Score 49.4; DB 17; Length 1512;
Best Local Similarity 47.8%; Pred. No. 0.00022;
Matches 143; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
Qy 132 CGGCTTCATCAAGCTCACCGACATGCGCGCTGGCGCGCCGAAAGGCTTCT 191

```

Db      669  CGGCTCGGCGAGGTGGGAGAGGTGCGACGGAGCGTCCGCGACGACCCCGGCAAGCT 728
Qy      192  CGAGAGCAGAGGCTGTGTGCACTGCAAGCGCGCACTGGAAGGTGTGATGA 251
Db      729  CGAGGCTGTGAGATGAGCGGCAAGACGGCATGTCTGTCACGAGAAAGCGGACCTCGA 788
Qy      252  TGGGCTGTGAATGAGGAACTGGAAGCTGTGCAATGTGGCGCGCGCTTACCGG 311
Db      789  CATGCCGTTTCAAGCGCGGTCTATGAGCTGTTCAGACACCGGCGAGGCTGTGTCTC 848
Qy      312  CAGCGTTGCTTGGCAGCAAGGCGCATATCGAGTGCCTTCAGCATGGCTTCAACG 371
Db      849  CAGCGAGCGGCTCATGTCCACAGAGACCTCTACGACGATTCAACAGCGTTCTGTCA 908
Qy      372  CAACGCGATTACGCTGTCCATGAATCTGGCATCGAGTAGAGCCGACATACCGCTG 430
Db      909  GATCGCGGAGTCGTTGCAGTGACGACCCCTTCGACGAGAGACGTTATGGGCGG 967

```

RESULT 10

```

US-10-282-122A-8343
; Sequence 8343, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8343
; LENGTH: 773
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-8343

```

```

Query Match      3.5%; Score 48.4; DB 17; Length 773;
Best Local Similarity 51.4%; Pred. No. 0.00035;
Matches 112; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```

```

Qy      658  CCGCGCGCGCAATGCGCTTCAAGCTGAGAGCGGACCATTTGCGTTATTCGCTGGC 717
Db      325  CCGCACCGATATATGTGTCAAGCTTTAGTAACATGATGATGATGTTTGTGTGGGA 384
Qy      718  GAGCCGTGGAACAGAGGCGCGTTCAGAGGCATAGCGTGCCTGTGATCAACCATGAA 777
Db      385  GAGCCTTGGAATATCTCAAGTGAACCTTTAGTTTAAAGTAAATGCTCTGCTCAATCAA 444
Qy      778  GAACCTGGAAGACACGCGGAAAAAGTCTTGCGCGTGAACCAACATAGGCGGAGAAA 837
Db      445  GACATATTTCCCAATGTGCAAGTAAAGTTCTGCGATTAATCAAGAAATGGGCTGAACAA 504
Qy      838  TACCCCAACCTATCTGGCGGAGACCAAGGGCGTAT 875
Db      505  CATCTCAACCTTATGTTGCTCCCTTACACGGCAATTAT 542

```

RESULT 11

```

US-10-369-493-43243
; Sequence 43243, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(5205)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 43243
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-10-369-493-43243

```

```

Query Match      3.4%; Score 48; DB 17; Length 744;
Best Local Similarity 47.9%; Pred. No. 0.00046;
Matches 138; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```

```

Qy      144  GCTCACCGACATGCGCGCTGCGGTGCGCCGCAAAAAGCTTCTTGAGAGCGAGG 203
Db      204  GCCGGGAGGATTGCGCGAGGTGGAAGTGTGTCAACAAAGGGGCTGGCGTGCGTT 263
Qy      204  CCTGTTCTGTGCACTGGAAGCGCAGCCCACTGGAAGGTGTGATGATAGGCTCTGAA 263
Db      264  GGAACCTGGCGCAGTGGGCCCGCGGTGAGAGACTGGGACACGATGTGGACACCAACGTGAA 323
Qy      264  TGGCGAATGAGCGGTGCGCATGCTGGCGCGGCGCGGTAGCGGCGAGGCTT 323
Db      324  GGGCTCTCTGTATGTGACCCACCGCGTCTGCCGATGTGGCGCGGACCGGAGCA 383
Qy      324  CGGACCAAGCGCGATATCGAGGTGCGTTTCAGATGGGCTTCAACGGGACCGATTTAC 383
Db      384  CGTGTCAACATGAGGGGTGCGTGGCGGAGTGGGCGGCTTACCGGGGCGCAAGGTGACGG 443
Qy      384  GGTTTCATGAATGATCTGGCATCAATGAAGCCGAACATACCGCTGGA 431
Db      444  CGCACCAAGGCGCTTGTGTGACACAGTTCAGCTCAACCTCCGCGGGA 491

```

RESULT 12

```

US-10-437-963-85118
; Sequence 85118, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

```

PRIOR APPLICATION NUMBER: 60/330 335
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/330 335

US-10-156-761-3146

Query Match	3.3%	Score 46.4	DB 15	Length 3048
Best Local Similarity	50.0%	Pred. No. 0.0025		
Matches 116, Conservative	0	Mismatches 116	Indels 0	Gaps 0

Qy 69 GGGTTTGCACATATGCCCGCATATGTCCGGGCACTTGGCAAGCTGGAAAAAGAAATTGAA 128
Db 645 GGTGCTGACGACCGCCCTGAAAGTCGTGTGTCCGCCCTTACACCGGCGAGAGAGACATCCC 704
Qy 129 ATTCCGCTTCATCAACGCTCAACGCAATGGCGCCGTGGCGGTGGCCGCGCAAAAAGGCTT 188
Db 705 GCTCGGCGATCCCATCTCGGCCGACGAGGCCGAACTCGAGGAGCGTGTGCGGCTGTT 764
Qy 189 CTTTCGAGAGACGAGGGGCTGTTGTCGCACTGGAAACGCGCAACTGGAAAGTGTCGAT 248
Db 765 CATCAACATGCTGATGTCGCGCACCGGATCTGTCCGCGACCCCACTTCGCGCGAGCTGAT 824
Qy 249 GGAATAGAGTCTGTAATATGGCGAACTGGACGGCTCGCACATGCTGGCGCCGCGCG 300
Db 825 CGAGCGGGTCCGCGACGCCCAATGGACGTGTTCACACCAACGAGACGTGCCG 876

RESULT 15

```

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTOKI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

```

Query Match	3.3%	Score	46.4	DB	15	Length	9025608
Best Local Similarity	50.0%	Pred.	No. 0.07				
Matches	116	Conservative	0	Mismatches	116	Indels	0
						Gaps	0

QY	69	GGGTTTACGATTTCCGCCGATGTGCGGGGACGTTGGCAAGCTGAAAAAGGAATTTGAA	128
Db	3936846	GGTCTTACGACGCGCCTGAAACGTGTGATGTGGCCGCTACACCGGCCAGAGAGACATCCC	3936787
QY	129	ATTGCGCTTACATCAAGCTCAACCGACATGCGCGCCTGGCGGTGGCCGCCGAAAAAGGCTT	188
Db	3936786	GCTCGGCATCCCATGTGCTCGGCGCGACGAGGCCGAACTGAGAGACGTGCTGGGCTGTT	3936727
QY	189	CTTCGAGAGCAGAGGGCCTGTTCGTGCACTTGAAGCGCAGCCCACTTGAAGTGTGTAT	248
Db	3936726	CATCAACATGTGTGTGTGTGCGCACCGAATGTGTCGGCGACCCCACTTTCGCCGAGCTGAT	3936667
QY	249	GGATAGGGTGTGTAATGGCGAACTGAGACGGCTCCCACTATGTGTGCGCCCGGG	300
Db	3936666	CGAGCGGGTGGCGGACCGCAACATGACAGGTGTACGACCAACCGAGACGTGTCCG	3936615

Search completed: May 20, 2005, 23:37:39
Job time : 899 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 20, 2005, 13:35:03 ; Search time 4942 Seconds
(without alignments)
10721.462 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaaacacatcatgacgc.....ccggcggaagtcgtggaat 1392

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.4	6.7	1872	9	CL982367 OsIFSC047
2	87.6	6.3	575	8	A2935061 BJ_Ba000
3	82.4	5.9	800	8	B2575941 msh2_470.
4	61.8	4.4	406	7	CP161028 B0692703-
5	58.6	4.2	1085	8	B2548675 pacsl-60
6	49.2	3.5	1035	8	B2563077 pacsl-164
7	48.6	3.5	1407	8	B2572739 msh2_2775
8	48.4	3.5	788	8	B2550725 pacsl-60
9	47.8	3.4	844	9	AL056552 Drosophila
10	47.2	3.4	459	6	CA718866 wkm2n.pk0
11	47.2	3.4	691	6	CD894397 G118_126C
12	47.2	3.4	824	7	CN129290 RH01_34
13	46	3.3	623	6	CA699640 w1k8.pk0
14	45.6	3.3	469	6	CA066446 SCROAD101
15	45.6	3.3	500	6	CA069366 SCSPAD106
16	45.6	3.3	602	6	CA069388 SCSPAD106
17	45.6	3.3	616	6	CA219519 SCSPAD106
18	45.6	3.3	645	6	CA212608 SCSPAD111
19	45.6	3.3	693	6	CA264998 SCSPAD111
20	45.6	3.3	708	6	CA069230 SCSPAD105
21	45.6	3.3	758	6	CA212728 SCSPAD112
22	45.6	3.3	840	6	CA174317 SCSPAD101
23	45.6	3.3	914	6	CA106423 SCQHR101
24	45.4	3.3	564	2	BE361762 DGI_82_E0

25	45.4	3.3	660	6	CD231149	CD231149 SS1_8_P06
26	45.4	3.3	663	6	CB925634	CB925634 ABAL_22_G
27	45.4	3.3	811	7	CN141741	CN141741 WOUNDI_1
28	45	3.2	503	8	CC351449	CC351449 OGIAE24TV
29	45	3.2	824	8	B2628435	B2628435 1h60b07.9
30	45	3.2	852	8	CC013120	CC013120 PUGHB08TB
31	44.8	3.2	568	2	BE775653	BE775653 MY-05-C-0
32	44.8	3.2	550	2	BE776036	BE776036 MY-10-B-0
33	44.6	3.2	472	4	BJ210804	BJ210804 BJ210804
34	44.6	3.2	580	4	BJ218083	BJ218083 BJ218083
35	44.6	3.2	604	6	CA253508	CA253508 SCJLFT10
36	44.6	3.2	686	6	CA134878	CA134878 SCJLFT10
37	44.6	3.2	695	9	CL164460	CL164460 104_358_1
38	44.6	3.2	942	8	B2575093	B2575093 msh2_430.
39	44.4	3.2	930	9	CL969992	CL969992 OsIFSC019
40	44.2	3.2	547	8	BH633034	BH633034 1007058E1
41	44.2	3.2	672	6	CA194351	CA194351 SCRFBI02
42	44.2	3.2	788	9	CC700945	CC700945 OGVBN08TH
43	44.2	3.2	1483	8	B2573056	B2573056 msh2_2923
44	44.2	3.2	1605	8	B2573144	B2573144 msh2_2963
45	44	3.2	424	9	CL794328	CL794328 OR_CBA000

ALIGNMENTS

RESULT 1
LOCUS CL982367 1872 bp DNA linear GSS 21-SEP-2004
DEFINITION OsIFSC047480 Oryza sativa Expressed library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL982367
VERSION CL982367.1 GI:52419216
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1872)
REFERENCE Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Zhao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
Unpublished (2004)
JOURNAL Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80486676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
location/Qualifiers
1..1872
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed library"
/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN
Query Match 6.7%; Score 93.4; DB 9; Length 1872;
Best Local Similarity 51.9%; Pred. No. 5.1e-14;
Matches 261; Conservative 0; Mismatches 236; Indels 6; Gaps 2;
QY CGCAGCGCCTTGTGCGTGACCCCGCCGCAATGCGTGCAGCGTGAAGCGG 692
DB 552 CGATGACGTGCGACGCTGTGTGCGCCCGCCGCGATGTATGAACTGGCGATTGG 611
QY 693 CACCATTTTCGTTATTCGTCGCGAGCGGTGAAACAGCAGGCGCGTTCAGGGCAT 752

Db 612 CAACATGATGCGGCTTTTGGCTCGGTGAGCGCGTGGAAAGCGAGGCGCATTAACGACCGCAT 671
 QY 753 AGCGCTGCGCGGTATCATCCGATGAAGACTCTGGAAGACACGCGGAAAAAGCTTTGCG 812
 Db 672 TGGCTTCACCGCTGCCACGTGCGATCTCTGGCGAGATCATCTGAAAAAAATTCGCG 731
 QY 813 CGTGACAAACAAATGGGCGGAGAAATATCCCAACACCTATCTGGCGGTGACCAAGGCGT 872
 Db 732 TACCGCTCGGACCTGGGTGAGAAAAATCCGACACCGCCGCGCTGGTGAAGCCGT 791
 QY 873 GATTCGGCGCGCATCTGCTGAGACGCGCAATTAACAAGAACCGCAAGACCATCGA 932
 Db 792 GATGGAAGCGCGCGCTGATGAGGCTGCCCGGAAAAACAAG---TGAAACCGCGCA 848
 QY 933 AATGCTGCGCAAAACAATAAGTGGTGTGACGTGAAAGTGTGGCGCTAGCATGAA 992
 Db 849 GATCTCTCGCGCGCGCGCTGCTCACTGCAAGAAACAGTATCTCACCGGCGGAATGCT 908
 QY 993 CGGCACTTGTGAATAGAAAAAGACATAAGCGCGCTACCGGACCTTCAACACTTTCT 1052
 Db 909 CGGGAATACGACACGCGCTGGGCGACGCGCTGACGAGACGCGACCCGATCCGTTCTT 968
 QY 1053 TCGCCACG---CGCGAGCTATCCGTCCTACAGACAGTGCAGTCTGATCTGACCGAGCT 1109
 Db 969 CAACGAGGCTGCGCTGATGATTCGTAACCTCTCTCGACGAGCATGCTTAAACCGATT 1028
 QY 1110 GAGCGCTGGGGCATGATCAATG 1132
 Db 1029 CCGCGCTGGGCGCTGCTTAAG 1051

RESULT 2
 A2935061 575 bp DNA linear GSS 24-APR-2001
 LOCUS B1_Ba0003E04r B. japonicum BAC library Bradyrhizobium japonicum
 DEFINITION genomic, genomic survey sequence.
 ACCESSION A2935061
 VERSION A2935061.1 GI:13777152
 KEYWORDS GSS.
 SOURCE Bradyrhizobium japonicum
 ORGANISM Bradyrhizobium japonicum
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Bradyrhizobiaceae; Bradyrhizobium.
 1 (bases 1 to 575)
 Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
 Golcochea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.,
 A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
 genome
 JOURNAL Genome Res. 11 (8), 1434-1440 (2001)
 MEDLINE 21376150
 PUBMED 11483585
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twing@clemson.edu
 Class: BAC ends
 High quality sequence stop: 535.
 Location/Qualifiers
 1..575
 /organism="Bradyrhizobium japonicum"
 /mol_type="genomic DNA"
 /strain="USD110"
 /db_xref="taxon:375"
 /lab_host="E. coli"
 /clone_lib="B. japonicum BAC library"
 /note="Vector: pindigo536; Site_1: HindIII"

ORIGIN
 Query Match 6.3%; Score 87.6; DB 8; Length 575;

Best Local Similarity 61.1%; Pred. No. 1.5e-12;
 Matches 160; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

QY 131 TCGCTTCATCAAGCTACGACATGCGCGCGTGGCGGCGCGGAAAAAGCTTCT 190
 Db 214 TCGGTTCAATTCGCTGATGATGCGCGCGCTGATCTCGCGCGTCAAGAGGCTTAA 273
 QY 191 TCGAGACGAGGCGCTTCTGTCGAATCTGAAGCGGCAAGCTCAATGGAAGTGTATG 250
 Db 274 CGCGCGCGAAGCGCTGACGTGAACCTGTGCGGAGGCTTCTGTGTCAAAGTCCGCG 333
 QY 251 ATAGGCTGTGAATGACGACATGACGCGCTGCACTGCTGGCGCGCGCGCTTACGCG 310
 Db 334 ACAAGCTCAATATCGGCTGTTGACGCGCCCATCTGCGCGCGCGCGATGCGCT 393
 QY 311 CCAGCGTTGGCTTTCG---CACCAAGCGCGATATGAGGTGCGCTTACAGATGGCTTCA 367
 Db 394 CTTCACTCGGCGCTGCGCGACGTCAAGGTCCGATGCGCGCGCTTCAATCTCGGATCA 453
 QY 368 ACGGCAACCGCATATACGCTGTC 389
 Db 454 ACGGCAACCGCATACAGGTGTC 475

RESULT 3
 B2575941/c 800 bp DNA linear GSS 17-DEC-2002
 LOCUS msh2_470.x2 msh Pseudomonas aeruginosa genomic clone msh2_470,
 DEFINITION genomic survey sequence.
 ACCESSION B2575941
 VERSION B2575941.1 GI:27211002
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 800)
 Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 Whole-genome-sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press
 COMMENT Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers
 1..800
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="M5H"
 /db_xref="taxon:287"
 /clone_lib="msh2_470"
 /note="Environmental isolate. Whole genomic shotgun
 library."

ORIGIN
 Query Match 5.9%; Score 82.4; DB 8; Length 800;
 Best Local Similarity 50.7%; Pred. No. 4.2e-11;
 Matches 225; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

QY 651 GGTGACCCCGCGCGCGCAATGCGTCAACGCTGGAAGCGGACCATTTTCGTTATG 710
 Db 526 GGTGTTTCACCGCGCGCGAGATGTCGCGCATCTCCAGGCGCGCATGACGCGGTTCTG 467
 QY 711 CGTGGCGAGCGCTGGAACGACGAGCGGCTTCAAGGCGATAGCGCGCGTATAC 770
 Db 466 CGCGCGGACCGCTTGGCGCGCGCTGCGTGAACGAGGCGGAGGCTTACCATGCGCAC 407

QY 771 CGATGAAGAACTCTGGAAGACACGCCGGAAGAAAGTCTTGGCGGTGACCAACAAATGGGC 830
 Db 406 CACCACGAGCATCTGGCGCGGACACCGGAAAGTCTCTGGGCACTACCGCGCTTCGT 347
 QY 831 GAGGAATATCCCAACACCTATCTGGCGGTGACCAACGCGCTGATTTGGGCGCGGANTCG 890
 Db 346 CGAGCGCTACCCCAACACCGCGCGGTGCTGATG3C--GCTCTCGACCGCGACCG 290
 QY 891 GCTGACGCGCGCAATTAACAAGACCGCAAGAACCCATCGAAATGCTGGCGGCAAAACA 950
 Db 289 CTTTATCGAACAAGACCGGCAAGACCGCTGGGACCGCGCACTATACGCGCGCGCA 230
 QY 951 ATACGCTGCTGACGCTGGAAGTCTGGCGGCTAGACATGACGCGCACTTTGCAATACGA 1010
 Db 229 CTATGTCGAGGCTCCGCTCGCGCGATCCAGCGCGCTTCTCGGCGCTACACGACGCG 170
 QY 1011 AAAAGACGATTAACGCGCGCTACCGGACTTCCACACTTCTTTCGCCACGCGCGCACTA 1070
 Db 169 CCTCGGCAACGCGCTGGACGACCGCATCCGCTGCTTATGCGACGCGCAATGCA 110
 QY 1071 TCCGCTCTACAGACGTCAGTCTG 1094
 Db 109 CCGGCGCTGCTCTCGAGGCGTG 86

RESULT 4

CP161028

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

CP161028 406 bp mRNA linear EST 25-JUL-2003
 B0692F03-5 NIA Mouse Blastocyst cDNA Library (Long) Mus musculus
 cDNA clone NIA:B0692F03 IMAGE:30457118 5', mRNA sequence.
 CP161028
 CP161028.1 GI:33270577
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 406)
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL MEDLINE 21429098
 PUBMED 11544199
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: B0692 row: F column: 03
 Seq primer: M13 Reverse
 High quality sequence stop: 406
 POLYA=No.

FEATURES

source

location/Qualifiers
 1..406
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="nia:EST:B0692F03-5"
 /db_xref="taxon:10090"
 /clone="NIA:B0692F03 IMAGE:30457118"
 /issue_type="Blastocyst"
 /dev_stage="3.5-dpc"
 /lab_host="DH10B"
 /clone_1lb="NIA Mouse Blastocyst cDNA Library (Long)"
 /note="Vector: pSPOR1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

extracted from a pool of 20 Blastocysts. Double-stranded
 cDNAs were synthesized with an oligo(dT) primer
 [Invitrogen:
 5'-pGACTAGTCTAGATCGGACGCGCGCCCTTTTCTTTT-3'] from
 0.2 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lone-linker L1-SalI, purified by phenol/chloroform, and
 separated from free linkers by centrifugation. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer SalI-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPOR1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.2 kb. The library was constructed
 by Yulan Piao (NIA)."

ORIGIN

Query Match 4.4%; Score 61.8; DB 7; Length 406;
 Best Local Similarity 52.5%; Pred. No. 1.5e-05;
 Matches 160; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 630 CGGCGAGAGCGCTTGTGTGCGTGAACCGCGCGCGCAATGCCCTCCACGCTGGAAGC 689
 Db 79 CGACACCGATGTCAGATCCGCGCAGTCCGCGCGCGGAAATGTTGCGAACCTGCGCGC 138
 QY 690 CGGACCATTTTGGTATTATGTCGCGCGGCGCGGCGTGAACACGAGCGCTGTTCAAGG 749
 Db 139 CGACACATGATGACGCGCTTCTCGCGCGCGCGCGGATGACACGCGCGGTCTATGACG 198
 QY 750 CATAGGCGTGCCTGATGATCAACGATGAAGAACTCTGGAAGACACGCGGAAAAAGTCTT 809
 Db 199 GGTGCGCTTCATCCATCTTGAACAAAGATATCTGGAAAGGCATCGCTGCGCGTT 258
 QY 810 CGGCGTGAACCAATGCGCGGCGGAAATACCCCAACACTATCTGCGCGTGAACGCG 869
 Db 259 CGGCGCTTCACAAAGAAATTCCTCAACACGCGTGGCAACCTATGCGCGCATTTCTGAATC 318
 QY 870 GCTGAT---TCGGGCGCGGATGCTGGCTGACGCGCGCAATTAACAAGACCGCAAGG 926
 Db 319 GATCATGATGACGACGCGCTTGGCGCAACGCGCGGACCGCAACGATTTGGCGAAGC 378
 QY 927 CATCG 931
 Db 379 AATCG 383

RESULT 5

BZ548675/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BZ548675 1085 bp DNA linear GSS 17-DEC-2002
 pac61-60_1334.61 pac61-60 Pseudomonas aeruginosa genomic clone
 pac61-60_1334, genomic survey sequence.
 BZ548675
 BZ548675.1 GI:27152256
 GSS.
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 1085)
 Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
 Burns, J.L., Kaul, R. and Olsen, M.V.
 Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press

JOURNAL MEDLINE 206857244
 PUBMED 11544199
 COMMENT Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu

FEATURES
source Class: shotgun.
Location/Qualifiers
1. .1085
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone_1ib="pac1-60_1134"
/note="clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN
Query Match 4.2%; Score 58.6; DB 8; Length 1085;
Best Local Similarity 52.2%; Pred. No. 0.00014;
Matches 153; Conservative 0; Mismatches 139; Indels 1; Gaps 1;

QY 84 GCCCGATGTCGGGGGCGGCAAGTGGAAAAAGAAATTGAAATTCGGCTTCATCA 143
DB 463 GGCCTGGGGCGGGCGGTTCCGACCGCAAGAAAGAGCGCCCTTGACATCGTTTCA 404
QY 144 GCTCACCGCATGCGCGCGCTGCGGTCGCGCGCAAAAAGCTTCTTCGAGGACGAG 203
DB 403 GGTACCGCATTCGGCTCGTATCGTGGCGCC -ACCATGCTTGGCCGACGCTACGG 345
QY 204 CCTGTTTCGTCACCTGGAAGCGGCGCACTGGAAGTGTGATGATGAGTGGTCTGA 263
DB 344 CTCACGCTCAATCTCAGGCGCCACCGTCTGCGGCACCTTGCAGCAAACTGCTCAG 285
QY 264 TGGCGAAGTGAAGCGGTCGACATGCTGGCGCGCGCGGCGTTGCGGCGACGTTGGCT 323
DB 284 CGCGGAGCTGATGATCGCGGATGCTGTACGGGCTGTCTACGGGCTCAGCTCGACCT 225
QY 324 CGGCACCAAGCGCGCATATGAGGTCGCTTCAAGCATGAGGCTTCAACGCGCAACG 376
DB 224 CGCGCGCAGGGCGGCAAGGAGATGGCGTGTCTCATGGACTGTGCAGAAAG 172

RESULT 6
BZ563077 1035 bp DNA linear GSS 17-DEC-2002
LOCUS pac82-164_4075.x1 pac82-164 Pseudomonas aeruginosa genomic clone
DEFINITION pac82-164_4075, genomic survey sequence.
ACCESSION BZ563077
VERSION BZ563077.1 GI:27186268
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1035)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Classes: shotgun.
Location/Qualifiers
1. .1035
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone_1ib="pac82-164_4075"
/clone_1ib="pac82-164"
/note="clinical isolate 2-164 Whole genomic shotgun"

FEATURES
source

ORIGIN library.
Query Match 3.5%; Score 49.2; DB 8; Length 1035;
Best Local Similarity 55.9%; Pred. No. 0.05;
Matches 113; Conservative 0; Mismatches 88; Indels 1; Gaps 1;

QY 658 CCGCGCCCGCAATGCGCTGACCGCTGAAAGCCGCGACCATTTTCGTTATTCGTCGGC 717
DB 448 CCAACCGCGCCAGATGTCGCGCATCTCCAGCGCGCGCGCATCGACCGGTTCTGCGCGG 507
QY 718 -GAGCGGTGAACACAGACGCGGTGTTCAAGGCGATAGGCGCGGTATCACCGATGA 776
DB 508 GGAACCTTGGGGCGCCCTGCGCGTGAACAGGCGCGAGGCTTCACCATGCGCACCGCA 567
QY 777 AGAAGCTGGAAGAGACACCGCGGAAAAAGTCTTCGCGCTGACCAAACTGGCGGAGAA 836
DB 568 GGGGATTTGGCCCGACACCGCGGAAAAAGGCGCTCGAATCAACCGCGCTTGTACCG 627
QY 837 ATACCCCAACCTATCTGCG 858
DB 628 CTACCCGAACACCGCGCGCG 649

RESULT 7
BZ572739 1407 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_2775.y2 msh Pseudomonas aeruginosa genomic clone msh2_2775,
genomic survey sequence.
ACCESSION BZ572739
VERSION BZ572739.1 GI:27207800
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1407)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Classes: shotgun.
Location/Qualifiers
1. .1407
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="msh"
/db_xref="taxon:287"
/clone_1ib="msh2_2775"
/clone_1ib="msh"
/note="Environmental isolate. Whole genomic shotgun library."

ORIGIN
Query Match 3.5%; Score 48.6; DB 8; Length 1407;
Best Local Similarity 50.7%; Pred. No. 0.078;
Matches 155; Conservative 0; Mismatches 148; Indels 3; Gaps 2;

QY 84 GCCCGATGTCGGGGCGGCAAGTGGAAAAAGAAATTGAAATTCGGCTTCATCA 143
DB 330 GGCCTGGGGCGGGCGGTTCCGACCGCAAGAAAGAGCGCCCTTGACATCGTTTCA 389
QY 144 GCTCACCGCATGCGCGCGCTG -GGGTCGCGCGCGCGGAAAAAGGCTTCTTCGAGGAGAG 202
DB 390 GCTGACCGATTCGCGCTCGTATGTGCGGCGCGCACCGAGGCGCTTGGCCAGCCTTACG 449

Oy		203	GCTGTTCGTCGAACCTGGAAAGCGCAGCCCAACTGGAAAGGTGTGATGATTAGCGTCTGA	262
Dd		450	GCTTCACGCTCAATTCTCAGAGCGGCCAACCGTCTTGAGCCACCCTTGCGGCACAAGTCTCTCA	509
Oy		263	ATGCGAATCTGACAGCGCTTCGACATGCTGCGCGCCGCGCCGTTTAGCGGCACGG--TTGG	320
Dd		510	GCGGCGAAGNGATATCGNNNGNAGTGCCTGTACGGGCGCTGATTACCGGGGTCAAGTCGG	569
Oy		321	CTTGGGACACCAAGCGCCGATATCGAGTGCCTTCACGATATGGGCTTCACAGCGCAAGCGCAT	380
Dd		570	CCTGCGCGCGCAGCGCGGACGAGCAAGGAGCGCGGTCTCTATGNACTGTGTCCACAAGCGCAC	629
Oy		381	TACGCT 386	
Dd		630	GGCGAT 635	
<hr/>				
RESULT 8 BZ550725/c BZ550725/c				
LOCUS DEFINITION				
BZ550725 788 bp DNA linear GSS 17-DEC-2007				
pacsl-60_2795.y2 pacsl-60 Pseudomonas aeruginosa genomic clone				
pacsl-60_2795, genomic survey sequence.				
ACCESSION BZ550725				
VERSION BZ550725.1 GI:27154306				
KEYWORDS GSS.				
SOURCE Pseudomonas aeruginosa				
ORGANISM Pseudomonas aeruginosa				
REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;				
Pseudomonadaceae; Pseudomonas.				
AUTHORS 1 (bases 1 to 788)				
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,				
Burns,J.L., Kaul,R. and Olsen,M.V.				
Whole-Genome-Sequence variation among multiple isolates of				
Pseudomonas aeruginosa library				
J. Bacteriol. (2002) In press				
CONTACT: Chris K. Raymond				
UNIVERSITY OF WASHINGTON				
Genome Center				
Box 352145, Seattle, WA 98105-2145, USA				
Tel.: 2062216954				
Fax: 2066857244				
Email: ckraymond@u.washington.edu				
Class: shotgun.				
<hr/>				
FEATURES				
source location/Qualifiers				
1..788				
/organism="Pseudomonas aeruginosa"				
/mol_type="genomic DNA"				
/strain="1-60"				
/db_xref="taxon:287"				
/clone="pacsl-60_2795"				
/clone_lib="pacsl-60"				
/note="clinical isolate 1-60 whole genomic shotgun				
library."				
<hr/>				
ORIGIN				
Query Match 3.5%; Score 48.4; DB 8; Length 788;				
Best Local Similarity 51.3%; Pred. No.0.078; Indels 3; Gaps 14;				
Matches 138; Conservative 0; Mismatches 128; Indels 3; Gaps 14;				
Oy		131	TCGGCTTCATCAAGCTCAACCGACATATGCGCGCTGCGCGCCGCCGAAAAAGGCTTCT	190
Dd		379	TCGGCTTACTGCGCGATCAACGCGCAAGCGCCATCTGTGGCCCATGCCAACGCGCTGT	320
Oy		191	TCGAGAACGAGGGCGCTTTCGTGCAACTGGAAGCCGACGCAACTGGAAGGTGTGATGG	250
Dd		319	TCGAGACCGAAGGCGATTCAGCCCGCAACGCGCGTCTTTGCGTAAGTGTGGCGCAGGTGA	260
Oy		251	ATAGAGTGTGAATGAGCGCAACTGACGCGCTCGACATATGCGCGCGCGCGCGTTAGAGCG	310
Dd		259	TCAAAGCGCTTATCTCCGCTCAGGTCAAC--GTATCCGCGCTGTGTGCGCCGATGCCCG	203
Oy		311	CGAGGTTGGCTTCGGCAACAGCGCGATATCGAGTGCCTTCAGCATGGGCTTCAAG	370

[illegible]

[illegible]

RESULT 10	
CA718866	
LOCUS	CA718866
DEFINITION	CA718866 459 bp mRNA linear EST 26-NOV-2006
ACCSSION	wk2m2.pK002.c3 wk2m2 Triticum aestivum cDNA clone wk2m2.pK002.c3 5'
VERSION	CA718866
KEYWORDS	CA718866.1 GI:25440659
SOURCE	EST.
ORGANISM	Triticum aestivum (bread wheat)
	Triticum aestivum

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 459)	Tingey, S. V., Powell, W., Wolters, P., Dolan, M., Hailey, C., Yuan, Z., Miao, G., Caraher, N. and Hanfey, M. K.	DuPont Wheat CDNA Sequence	Unpublished (2002)	Contact: Scott V. Tingey Crop Genetics E. I. DuPont de Nemours and Company 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA Tel: 302-631-2602 Fax: 302-631-2607 Email: Scott.V.Tingey@usa.dupont.com Seq primer: MJ3.

FEATURES	SOURCE
Location/Qualifiers	1. .459
	/organism="Triticum aestivum"
	/mol_type="mRNA"
	/cultivar="Spring wheat"
	/db_xref="taxon:4565"
	/clone="wkm2n.pk002.c3"
	/tissue_type="kernel"
	/lab_host="DH10B"
	/clone_id="wkm2n"
	/note="vector: pBluescript SK+, Site_1: EcoRI, Site_2: XhoI; wheat (Triticum aestivum L.) kernel malted 175 hours at 4 C, normalized"
ORIGIN	

	Query Match	3.4%	Score 47.2	DB 6	Length 459
	Best Local Similarity	50.9%	Pred. No. 0.15		
	Matches 112	Conservative	0	Mismatches 108	Indels 0
QY	156	GCGCGCGCTGGCGGTGGCCGCCGAAAAAGGCTTCTTCAGGACGAGGGCTGTTCTGCA	215		
DB	48	GCGAGCATATGCGACGCGCAGCAGGTGTGGACAAACGGGACGAGGTGGGCCCGACCGGGGA	107		
QY	216	ACTGGAAGCGCAGGCGCAACTGGAAAGTGTATGATGATGGGTCTGTAATGGCGAACTGGA	275		
DB	108	CTCGGGGGCGTTTGGCTTTCGCCGCTGTAAGCGAGGAGGGCGCCGAGTGTGAGGCCGG	167		
QY	276	CGGCTGCACATGCTGCGCGCCGCGCCGCTTAAAGCGACGCGTTGGCTTTCGACCAACAGGC	335		
DB	168	CGGCTTCGGGCGCAGCGCGTCCGAGTCCCTCGAGACGAGGGTTCGACGACGCGCAACATCA	227		
QY	336	CGATATGAGAGTGCCGCTTCAGCATATGGGCTTCAACGGCAAC	375		
DB	228	GGAAGCGCAGTCTCTCGCTCGCGAGGGCTCTTCTCTCAAC	267		

RESULT 11
CD894397

LOCUS	CB894397	691 bp	mRNA	linear	EST 14-JUL-2003
DEFINITION	G118.126C2F010823	G118	Triticum aestivum	CDNA clone G118126C22,	
ACCESSION	CB894397				
VERSION	CB894397.1	GI:32665884			
KEYWORDS	EST.				
SOURCE	Triticum aestivum (bread wheat)				
ORGANISM	Triticum aestivum				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 691)	Genopiante.	Genopiante, a major partnership french program in plant genomics	Unpublished (2003)	Contact: Genopiante

207 rue Henri-Rochefort 91023 Evry-Corbeil France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).

```

FEATURES
source
location/Qualifiers
1..691
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="rectal"
/db_xref="taxon:4565"
/clone="G118126C2"
/tissue_type="grain (118 degrees per day after
pollination)"
/clone_lib="G118"

```

Query Match	Best Local Similarity	3.4%;	Score 47.2;	DB 6;	Length 691;
Matches 112;	Conservative	50.9%;	Pred. No. 0.16;	Mismatches 108;	Indels 0;
					Gaps 0;
QY	156	GCGCCCGCTGCGCGGTGCGCCCGCAAAAAGGCTTCTTCGAGGACGAGGGCCCTGTTCTGTCGA	215		
Db	171	GCGCGCAATGGCAGCGCGCACAGAGTCGCGCAACGCGAGCGAGGTGGGCCCGACCGGGGGA	230		
QY	216	ACTGGAAAGCCGACGGCCAACTGGAAAGTGCGATGATGAGGTGCTGTAATGGCGCAACTGGA	275		
Db	231	CTCGGGGGGCGTTGCTTCCGCGCGGTGAAGCGGAGGGAAGGGCGCCGGAGTTCGAGGCCGCG	290		
QY	276	CGGCTCGCACATGCTGCGCGCGCGCGCTTAGCGGGCAGCGTGTGGCTTTCGGCACCAAGGC	335		
Db	291	CGGCTCGCGCGGACGCGCGTCCGAGTCTTCGAGAACGAGGGTTCGACGACGCGCAACATCCA	350		
QY	336	CGATATCGAAGGTGCGCTTCAGCATTTGGGCTTCAACCGGAAC	375		
Db	351	GGAAGCCGAGTCTCTCGCTCCGCGAAGGGCTCTTCCCTTAAC	390		

RESULT_12	
CN129290	
LOCUS	
DEFINITION	CN129290 824 bp mRNA linear EST 01-APR-2004
ACCESSION	RHOH1_34_C09.g1_A002 Acid- and alkaline-treated roots Sorghum
VERSION	bicolor cDNA clone RHOH1_34_C09_A002 5', mRNA sequence.
KEYWORDS	CN129290.1 GI:45956911
SOURCE	EST.
ORGANISM	Sorghum bicolor (sorghum)
	Sorghum bicolor

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 824)	Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Lim, S., Eastman, A. and Pratt, L.H.	An EST database from Sorghum: acid- and alkaline-treated roots

Query Match 100.0%; Score 2443; DB 8; Length 464;
 Best Local Similarity 100.0%; Pred. No. 1e-217;
 Matches 464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIIRSSSKKLLTLTASIAVWGLTIAPDVAGVGLKEKEDLKFGFIKLTDMAPLVAAB 60
 DB 1 MKTIIRSSSKKLLTLTASIAVWGLTIAPDVAGVGLKEKEDLKFGFIKLTDMAPLVAAB 60
 QY 61 KGFEEDEGLFVLEAQAANKVMDRVVNGELDGSNMLAPAPLAASVFGTKADIEVPSM 120
 DB 61 KGFEEDEGLFVLEAQAANKVMDRVVNGELDGSNMLAPAPLAASVFGTKADIEVPSM 120
 QY 121 GFGNNAITVSNELIWHQMKPNIPLEGKPVHPRIKADYLKPVVEKYKAEKPPNMAATPPAG 180
 DB 121 GFGNNAITVSNELIWHQMKPNIPLEGKPVHPRIKADYLKPVVEKYKAEKPPNMAATPPAG 180
 QY 181 SHNIKLRVWLAAGINPGYSPPODISGOIGADALLSVTPPPQMPSTLEAGTIFGYCGB 240
 DB 181 SHNIKLRVWLAAGINPGYSPPODISGOIGADALLSVTPPPQMPSTLEAGTIFGYCGB 240
 QY 241 PMNQAVFGIGVPIITDELMKDPKVPFGVTKQNAEKYPNTYLAATKALIRAAIWLDA 300
 DB 241 PMNQAVFGIGVPIITDELMKDPKVPFGVTKQNAEKYPNTYLAATKALIRAAIWLDA 300
 QY 301 DNNKRRKEAIEMLAQKQYGVADVEVLAASNGTEFEYKDKRALPDPTFFRHGASPSY 360
 DB 301 DNNKRRKEAIEMLAQKQYGVADVEVLAASNGTEFEYKDKRALPDPTFFRHGASPSY 360
 QY 361 SSAVWVLTQLRRWGMINEFKPDNMYLDTAKVYRPDIYLAALKEVLAEGKAKAEDPPADT 420
 DB 361 SSAVWVLTQLRRWGMINEFKPDNMYLDTAKVYRPDIYLAALKEVLAEGKAKAEDPPADT 420
 QY 421 STKPSQNFIDKVPDPANKPNDYLAFAIGLKGKQTVAGGKVD 464
 DB 421 STKPSQNFIDKVPDPANKPNDYLAFAIGLKGKQTVAGGKVD 464

RESULT 2
 ABO63846.
 ID ABO63846 standard; protein; 456 AA.

AC ABO63846;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seqid 10363.
 XX
 KW Recombinant expression vector; transcription regulatory element;
 XX Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 PN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M,
 XX
 DR WPI; 2003-895346/82.
 DR N-PSDB; ACH97397.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 XX preparing a vaccine composition against Klebsiella pneumoniae.
 PS Disclosure; SEQ ID NO 10363; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella

CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

SQ Sequence 456 AA;

Query Match 23.6%; Score 577.5; DB 7; Length 456;
 Best Local Similarity 33.5%; Pred. No. 2.2e-44;
 Matches 153; Conservative 76; Mismatches 173; Indels 55; Gaps 17;

QY 7 SSSKLLTLTASIAVWGLTIAPDV-----GAVGLKEKEDLKFGFIKLTDMAPLVAAB 60
 DB 44 SISRRRLQAGALG--GAMLLPGVQAAMAGSDPBEQTRVAVGPIPLTDCAPLAIATA 101
 QY 61 KGFEEDEGLFVLEAQAANKVMDRVVNGELDGSNMLAPAPLAASVFGTKADIEVP-PS 119
 DB 102 KGFDOQRYGITVLAASKASMAAVRDKIVAGELDAHILYGLYGLGLIASKPOAMANLMT 161
 QY 120 MGFNGNAITVSNELIWHQMKPNIPLEGKPVHPRIKADYLKPVVEKYKAEKPPNMAATPPA 179
 DB 162 LNRNGAAILTSEL--QEKGVTDLG-----LKRLLDR-SAPGS-YTFATTPPT 206
 QY 180 GSHNIKLRVWLAAGINPGYSPPODISGOIGADALLSVTPPPQMPSTLEAGTIFGYCGB 239
 DB 207 GTHAMWLYYWLASAGIDP-----FNDVATVVVPPQVMNMRIGMSGFCYG 253
 QY 240 EPMNQAVFGIGVPIITDELMKDPKVPFGVTKQNAEKYPNTYLAATKALIRAAIWLDA 299
 DB 254 EPMNARAIANDRIQFTNAITSODIWIPEHPEKVLGTRRWVERNPVTARALVAALAEQRWI- 312
 QY 300 ADNNKRRKEAIEMLAQKQYGVADVEVLAASNGTEFEYKDKRALPDFTN--TFFRHG-AS 356
 DB 313 AASPEMTRETAARLARGMILNTRKEQYLTRMLG--EYDGLGRMQDAPRIRPWAGGEVS 370
 QY 357 YPSYSSAVVLTQLRRWGMINEFKPDNMYLDTAKVYRPDIYLAALKEVLAEGKAKAEDP 416
 DB 371 FFWLSDGMWFLTQFRWGLLQKQ-APD--YLAVASRINRIDVQAQAQ--AVGGISA--- 421
 QY 417 PADTSIKPSQNFIDKVPDPANKPNDYLAFAIGLKG 453
 DB 422 PA--ARNRSTILMDGTVMGSDPEGYARHFSIQKKG 455

RESULT 3
 ABU42051
 ID ABU42051 standard; protein; 403 AA.

AC ABU42051;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #27578.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Pseudomonas syringae.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002W0-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.

Pt Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zykkind JM,
Xi Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
Xr WPI; 2003-029926/02.
Xs N-PSDB; ACA45921.
XX
Ys New antisense nucleic acid, useful for identifying proteins or screening
Yx for homologous nucleic acids required for cellular proliferation or
Zt isolate candidate molecules for rational drug discovery programs.
Zx
AA Claim 25; SEQ ID NO 69975; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs; or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences

Sequence 403 AA;

Query	March	21.5%:	Score 525;	DB 6:	Length 403;
Match	Local Similarity	31.7%:	Pred. No. 1.4e-39;		
Base	132;	Conservative	72;	Mismatches 164;	Indels 48; Gaps 12;
Qy	38	EKEDLKPFPIKLTDMAPLVAABAKKFFPEDEGLFVQLAEQANMKVVMRDVRVNGELDGSHTL	97		
Db	19	EMNSLDVGFMALTDCAPLVVAATGQFAQPYGSLNLKXQTSWAGLRDRLVSQGLQAASLT	78		
Qy	98	APAPLAAVSGF--GTKADIEVPFSKGFEGNAILYTSNEIMWQMKNPILBEGSKVPHPIAAD	155		
Db	79	YGLIYAVELGISGGPATDMALIMGINQOCINLSRE-----LQDAGVITPBEALD	128		
Qy	156	YLKPIVREKYKAEKPEFNMAATFPAGSHNIKIRYMLAAGINPGYSSPQDISQIGADAL	215		
Db	129	-----KRAHSGSGSLTFRQTFPTFCNHNAMMLTYMLASGIIIP-----LDVV	168		
Qy	216	LS-VTPPQPMSTLEAGTIFGVCGEPPNQQAIVEKIGCVPIITBEIMKDTPEKVFVTK	274		
Db	169	TSVVVPPFQMAQHQAAGRIDGFCVGEPPGASAVOQOLGFTWATSGAIPMDHPGKVLGCTR	228		
Qy	275	QMAEKYPTLYAVLKALIRAAIMLWDADNNKQRKKAIEMLAQOYGAIVETLAASMNQTF	334		
Db	229	EFVEONPITARALIMAVLEASRFIE-QSDHNRKSTQAQLSGVDYIDASLDLCIEPRLAQY	287		
Qy	335	EYEKDQRALPDPFTNFFPHG--ASYPSSASAWYLTQLRRGMVINEFKEDNMYLDTAKXVY	393		

[illegible]

XX	RESULT 4
XX	ADL05206
XX	ADL05206 standard; protein; 474 AA.
XX	
AC	ADL05206;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	M. catarrhalis protein #972.
XX	
KM	Moraxella catarrhalis; infection.
XX	
OS	Moraxella catarrhalis.
XX	
PN	US673910-B1.
XX	
PD	06-JAN-2004.
XX	
PF	04-APR-2000; 2000US-00540236.
XX	
PR	08-APR-1999; 99US-0128416P.
XX	
PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	
PI	Breton GL;
XX	
DR	WPI; 2004-178127/17.
DR	N-PSDB; ADL03286.
XX	
PT	New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT	preparing a composition for diagnosing, preventing or treating infection
XX	caused by Moraxella catarrhalis.
XX	
PS	Disclosure; SBQ ID NO 2892; 429pp; English.
XX	
CC	The invention relates to an isolated nucleic acid encoding an Moraxella
CC	catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC	composition for diagnosing, preventing or treating infection caused by
CC	Moraxella catarrhalis. The present sequence represents the amino acid
CC	sequence of a M. catarrhalis protein.
XX	
SQ	Sequence 474 AA;

Query Match	18.3%	Score 446.5	DB 8	Length 474
Best local Similarity	27.9%	Pred. No. 3,4e-32		
Matches	129	Conservative 82	Mismatches 180	Indels 71
Gaps	11			
Qy	2	KTIIRSSSKLLLTLSASLAWGLTIADVGAVGKL--EKEDLKFGFIKLTDMAPLAVAA	59	
Db	79	KTVGKTTA-----LAATASVLPATLTQEAALITLKEKKSVDIGFPILCATPLIMAD	132	
Qy	60	EKGFFEDGFLVOLAQNMKVMDRVYNGEIDGSHMLAPPLIASVGGT-KALIEVPF	118	
Db	133	PLGYAEGSICANILTKRAGMALVRDQNNRELDAAHFLAPPIALINLGLGSKQNMKVA	192	
Qy	119	SMGNGNAITVSNBIWHQMKENIPLEGKPVHPIKADYLKPFVEKKYKAEKFNMAATPT	178	
Db	193	IQNTNGALVAA-----LKKNNNNPKNQMGMPAIPFE	226	
Qy	179	AGSHNIKLRMYLAAGINPGYRSPQDISQIGADALSTVPRPOMPSTLEGGTFGCV	238	
Db	227	HSIRHYILRTYLAHGGDP-----DKDVKRLTTPPMINLVAKNIDGCFG	273	
Qy	239	GEPNNQAVFGKIGVPIVTDDELMKDPKEKVFYTKQMAEKYPNTYLAATVALLPAALWL	298	
Db	274	PEPNQKRVAMDKAGIHTLSRDINDNGHCCGCFGSGISGINDYPOFLMALYALIKANVM	332	

CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 249 AA;

Query Match 12.0%; Score 292.5; DB 7; Length 249;

Best Local Similarity 31.3%; Pred. No. 2.7e-18; Matches 73; Conservative 40; Mismatches 105; Indels 15; Gaps 6;

QY 218 VTPPQWPTSLAGTIFGVCGEPMNOQAVFKIGVITIDELMKDPEKVFSTQMA 277
DB 19 VVEPAGWGLQGRIDGFCAGGPGWGLAVDQGGFTIATSOAWPHEKVLGTAAV 78
QY 278 EKTPNTYLAATKALIRAIWLDNKNRKEALIMLAOKOYGVADVLAASNGTEYE 337
DB 79 DAVPNTARALVMAVLDASRFE-ONMENRGTALQISGRVYDAPLAIQPRFGRYODG 137
QY 338 KDKRALPDNTFFRHG-ASYPSVSAVWYLTQRRKGMINEFPDQWYLDLKNVVRPD 396
DB 138 LGNAMODPPLRFRYADSEVNRPMUSDGMMWMTQFRKGLRE-DPD--YLGIAARRVQOTA 194
QY 397 IYLAAKELVAEGKAKADEFPADTSIKPSQNFIDKVPFANKENDYLAFAI 449
DB 195 LYDAAATFALGL--RLDGDWRRST-----LIDGRTWGSDPDAGYARSPFI 237

RESULT 7

ABU38446
ID ABU38446 standard; protein; 399 AA.

XX ABU38446;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #23973.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pseudomonas aeruginosa.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haeselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Treweek JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA42316.

XX Claim 25, SEQ ID NO 66370; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 399 AA;

Query Match 9.8%; Score 240; DB 6; Length 399;

Best Local Similarity 24.6%; Pred. No. 4.1e-13; Matches 93; Conservative 68; Mismatches 141; Indels 76; Gaps 14;

QY 9 SKLLTLTSLASLAWG-LTIAPDVAGVGLKEB-DLKFGRFKLTDMAPLVAANEKGFED 66
DB 9 SRBDILKALTLAAGALPLLSLQAPAAAPAPRIGVLPITDAPPLVAHANGFEA 68
QY 67 EGIPLVLEAQA---NMKVMDRVNNGELDGSNMLAPPLAASVGFKALIEVPSFGFN 123
DB 69 EG--IQAEFVLLRSMQVTEAFISGVNVTHLSPTWVAR--YGSKVPAKVVAWNVG 124
QY 124 GNAITVSNELWHQMKNPILPEGKPVPIKADVLKPVVEKYKAKGFNNAMTFPAGSHN 183
DB 125 GSGITVAPET---ADVRLGKSV-----AIPFWSTHN 155
QY 184 IKRLWYLAAGINPGYSPDISGOIGA---DALLSTVPPQWPTSLAGTIFGVCV 238
DB 156 VVLQQLRONGL-----RAVSRAVGALLADENVVLVLPSPDMPALASKRIHYIV 207
QY 239 GEPWNOQAVFKIGVPTIDELMKOTPEKVFSTQMAEKYNTYLAATKALIRAIWL 298
DB 208 AEPFNALAEMLKVGRTVQRTGDMWRHACCVFMHEHDLERRRPOWSQKVNAIVKQLW- 266
QY 299 DADNNNRKEALIMLAQ---KQYVGADVETL-----AASNGTFEYEXKDKRA 343
DB 267 ---TTRHRAEAQLSLAGANRRITPHAPVGLAVLAPGAEQQAAYLASGAIIRHADMORR 323
QY 344 LPDPNTFFRHGASYPYS 361
DB 324 I-DFQPY-----PYPSYT 335

RESULT 8

ADSI4947
ID ADSI4947 standard; protein; 399 AA.

XX ADSI4947;

XX 16-DEC-2004 (first entry)

XX Pseudomonas aeruginosa quorum sensing controlled protein, SEQ ID 502.

XX Cytostatic; Anti-HIV; Antimicrobial; Immunosuppressive;
XX quorum sensing signaling; bacterium; quorum sensing controlled gene;
KW biofilm-associated disease; cystic fibrosis; AIDS; middle ear infection;

[illegible]

Query Match	9.8%	Score 240;	DB 7;	Length 419;
Beet Local Similarity	24.6%	Pred. No. 4.4e-13;		
Matches	93;	Conservative 68;	Mismatches 14;	Indels 76; Gaps 14
Db	9	SKILLITLSASIAWVG-LTIAPDVGAVKLEKE-DLKGFIKLTQMAPLAVAAEKGFED	66	
Db	29	SRRIILKIALALISAAGLPLISSIQARAAAPDPAVRGIVLPITDPTLVVHANGLEA	88	
Qy	67	EGLFVQLEAQA--NMKVYMDRVVNGELDGSHMLAPAPLAASVGFGRYADIEVPSMGFN	123	
Db	89	EG--IQARPPYILSSMAQVIAFTISGQVNVVHLLSPMTVMAR--YGSKVPYAVMNHVG	144	
Qy	124	GNALTYSNEITHQKKPNIPLEGGRPYPIKADYLKPVVEKTKAEKGPNNMAMTFPAGSIN	183	
Db	145	GSGGLTVAPEI---ADVROGRGKSV-----AIPFWYSIHN	175	

QY 184 IKRLVLAAGINPGYSPDISGQIGADALSTPPMPSTIEAGTIGCY 238
 DB 176 VILQQLRLNGT-----RAVSRVAGALAADEVNLVLPSPDPPALAKRIRHGY 227
 QY 239 GEPNNOAVFGKIGVPTIDBELMKDPEKVGVTWKMAEKYPTVLAVTKALIRAIWL 238
 DB 228 AEFNALAEVTKKGRQRFQGTGVDVWRNACCVFMHEHDLERRRQMSKVNAIVKQALM- 286
 QY 299 DADNNKRRKAEIMLAQ---KQYVADVEVL-----AASNGTFEYKDKRA 343
 DB 287 ---TREHRAAQAQLLSKAGANRTPHAPVTLGRTVLAAGAEQQAYLASGAIIRADQERR 343
 QY 344 LPPDPTFFRHGASYPSTYS 361
 DB 344 I-DFOPY-----PYPSYT 355

RESULT 10

ABU21883
 ID ABU21883 standard; protein; 396 AA.

AC ABU21883;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #7410.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Burkholderia fungorum.

PN W0200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELITR-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Treweek JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX MPI; 2003-029926/02.

XX N-PSDB; ACA25753.

XX Claim 25; SEQ ID NO 49807; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene in which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 396 AA:

Query Match 9.6%; Score 234.5; DB 6; Length 396;
 Best Local Similarity 23.3%; Pred. No. 1.3e-12;
 Matches 105; Conservative 77; Mismatches 150; Indels 119; Gaps 18;

QY 9 SKKLTLTSLASLAWGTLTADVGAVGKLEKED---LKRFGIKLDMAPLAAEKGFF 64
 DB 7 SRREWLKLSMFTVAG--AAPLIALNARAENRDPVRIGYLPITDAAPLVAHNNGYF 64
 QY 65 EDEGLFVQLEA--QANKVVDADVNGELDGSNMLAPAPLAASYFGTKADIEVFFSMGFN 123
 DB 65 DASGLAVEKPTLLRSNAQLVEAFLSQGVNVVHLAPRTIAR--YSGQAPAKVAVANHN 122
 QY 124 GNAITVSNEIWHQMKPNIPLEGKPVNPIKADYLKPEVEKYKABGKPFNMAITPPAGSHN 183
 DB 123 GSALTVAAPDI-----GKLGEIGGKTV-----AVPFWSIHN 153
 QY 184 IKRLVLAAGINPGYSPDISGQIGADALSTPPMPSTIEAGTIGCYVCEBPW 242
 DB 154 VVVQHMLRAGL---VPLERKDELKANEVLLVWSPDMPALASROIAGIYAEPP 208
 QY 243 NQOAVFGKIGVPTIDBELMKD-----TREKVFQVTWKMAEKYPTVLAVTKALIRAA 295
 DB 209 NAAAEELKVGKVLRFQGDVWKNACCVFMHEHDLERBAWMSK-----VVDAAVKAQ 261
 QY 296 IWLADADNNKRRKAEIMLA---OKQYVADVEVLAAASNGTFEYKDKKALPDPTFFR 352
 DB 262 VMTTRA-----HPQAAQLLSKSGRNHTPHSANVLT---NVLAPPGDEGRVLAD----- 308
 QY 353 HGASYSYSAAVYTLQLRKMGWINEFKPDMWLDTRAKN-VRPDIYLAARELV----- 406
 DB 309 ---RAITHADWH-----AKRIDFQPYPPAYATELVRRLKA 341
 QY 407 --AEGKAK-----AEDFPADTSIKPS 425
 DB 342 TQVEGNHQPLQHDPAFVARDLVDRPVKKS 372

RESULT 11

ABU16603
 ID ABU16603 standard; protein; 257 AA.

AC ABU16603;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #2130.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS *Acinetobacter baumannii*.

XX W0200277183-A2.

XX 03-OCT-2002.

XX

PF 21-MAR-2002; 2002MO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haeelbeck R, Ohlsen KU, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
 XX
 DR MPI: 2003-029926/02.
 DR N-PSDB; ACRA20473.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 44527; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained in
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SO Sequence 257 AA;

Query Match 9.3%; Score 227; DB 6; Length 257;
 Best Local Similarity 23.5%; Pred. No. 3.4e-12;
 Matches 69; Conservative 52; Mismatches 122; Indels 50; Gaps 7;

QY 91 LGGSHMLAPAPLAASVGFGRKAD-----TEVPMSMFPNGATVSEIWHQM-----KPMI 141
 DB 1 LPAHGXSMPLPAAANG-----ADQIGIALQTPVLVSKNNAFISLSQKLIYQLAIGESDA 56
 QY 142 PLEGKPVHPIKADYLUKPVVEKYEKGFNNMAMTTPAGSHNIKLYWTLAAGINPGYS 201
 DB 57 QTTAAQVIVYIGDH-----TUSLAHVYKSHIHVYCLAEWALA----- 95
 QY 202 PPODISGQIGADALLSVTPPQMSLTLEAGTIFGCVGSEPMNQAVFKSIGVPVITDEBL 261
 DB 96 -----DSRIAQTLKIKALPPPYVVEALDHWVIDGFCVGEPPMNTQGLIGLSKIVCSQDI 150
 QY 262 WKDTPKVGKVKQMAEKYPTNYLATATKALIRALVLDADNNKRNREALEMLAQKYVGA 321

DB 151 IPVNAKVALAVTQEMABQHPQTLVATLTKAQ--KEISNLDPAPILKTL----- 200
 QY 322 DVEVLAASWNGTFEVEKDKRALPDNTEFRHNGASYPSSAVWYLTOLRRWG 374
 DB 201 -VEFGIVRFHSCSEVAVDKRYMIGNIVKLVENANAPQEDDFHMLQGMQKMG 252

RESULT 12

AAG91152
 ID AAG91152 standard; protein; 294 AA.

XX
 AC AAG91152;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 4906.

KW Corynebacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

PD 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99QP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayaashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR N-PSDB; AAH66371.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying

PT expression point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

PS Claim 17; SEQ ID NO 4906; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of Corynebacterium bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Corynebacterium bacterium, and identifying a homologue of a gene derived from

CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the European Patent Office

SO Sequence 294 AA;

Query Match 9.3%; Score 226; DB 4; Length 294;
 Best Local Similarity 23.6%; Pred. No. 5.2e-12;
 Matches 78; Conservative 56; Mismatches 152; Indels 44; Gaps 6;

QY 1 MTTIRSSSKULLLTLSASLAWGLTAPDVCAVGLKEKD--LKFGPIKLTDMAPLAV 57
 DB 1 MTHILFDSRRFLQAFASLSTALGAARYVSTSNNEPADTPTLTIGVPIAGSPAI 60
 QY 58 AAKEGFEEDEHGFVQLEQANKKVMVDRVWVNEGLDGSHTLAPAPLAASVGF--GTAKDI 116
 DB 61 ADALGLFKKGGVNVTLKKTSGMSDILWTATVATQQLDVAHMLSPWYAINAGVTNARPT 120

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0242923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haeelbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI; 2003-029926/02.
DR N-PSDB; ACA25903.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 49957; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIFO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 391 AA;

Query Match 9.0%; Score 220.5; DB 6; Length 391;
Best Local Similarity 24.4%; Pred. No. 2.6e-11;
Matches 94; Conservative 63; Mismatches 136; Indels 93; Gaps 16;

QY 9 SKKLLTTSASLAWGLTAPDVAGVKEKED---LKFPIKLTDMAPLVAAREGFF 64
DB 2 SREWMKLKSMFVAG--AGPLLGALNABAAPDPNPAVIGYIPIDANPLVAHNGGF 59
QY 65 EDEGLFVQLEA-QAMNKVMDRVNNGELGSHMLAPAPLAASVGFCTKADIEVPSMGFN 123
DB 60 EAAAGLAVERPTLLRSWAQLVEAFVLSQVWVHLLAEMTWAR--YGSQAPAKVAVANNHN 117
QY 124 GNAITVSNIEHWQMKNIPLBEGSKPHPIKADYLKVEKYEKGFPMAMFPAGSHN 183
DB 118 GSALTIVSPIT---GKVADLGGKIV-----AVPEWSTHN 148
QY 184 IKLRWYLAAGINP-----GYSSPPODISQIGADALSVTPPPQMPSTLEAGTIPGYCV 238
DB 149 VVVQDMLRTQGLVPLVKDGLKP-----NEVRLTVMAPSDMPALASQINGFIV 199

QY 239 GERHNOQAFKGIQGVITDEELMD-----TPEKVGCVTKQNAEKYPNTYLAATKL 291
DB 200 AEFNNAAGELKIGKALFTGVMKNHACCVFMHERDLTERAAMSQK-----VVDV 252
QY 292 IRAATWLDADNNKRNKEATLEMLA---QKQY-----VGADVEVLAASWNGTFEYKDDKRAL 344
DB 253 VKRQGLMTRA---HQRBAAGLSKQGNHTHTTAVALTEVLAPAGDGRYAD--RAI 306
QY 345 -----PDFNTEFRHGAASYPSYS 361
DB 307 IHADWIAKRIDFQPY-----PYPAYT 327

RESULT 15

AAB76701
ID AAB76701 standard; protein; 275 AA.

XX
AC AAB76701;XX
DT 11-APR-2001 (first entry)XX
DE Corynebacterium glutamicum MCT protein SEQ ID NO:384.XX
XX Corynebacterium glutamicum; brevisbacterium lactofermentum; MCT;XX
XX membrane construction and membrane transport protein; petroleum spill;XX
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;XX
XX identification; microorganism; fine chemical production; transformation;XX
XX genome mapping; genetic engineering.XX
XX Corynebacterium glutamicum.XX
PN WO200100805-A2.XX
PD 04-JAN-2001.XX
PF 23-JUN-2000; 2000MO-1B000926.

PR 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031454.
PR 08-JUL-1999; 99DE-01031478.
PR 08-JUL-1999; 99DE-01031563.
PR 09-JUL-1999; 99DE-01032122.
PR 09-JUL-1999; 99DE-01032124.
PR 09-JUL-1999; 99DE-01032125.
PR 09-JUL-1999; 99DE-01032128.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032182.
PR 09-JUL-1999; 99DE-01032190.
PR 09-JUL-1999; 99DE-01032191.
PR 09-JUL-1999; 99DE-01032209.
PR 09-JUL-1999; 99DE-01032212.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032228.
PR 09-JUL-1999; 99DE-01032229.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032927.
PR 14-JUL-1999; 99DE-01033006.
PR 14-JUL-1999; 99DE-01033006.
PR 14-JUL-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040830.
PR 27-AUG-1999; 99DE-01040831.
PR 27-AUG-1999; 99DE-01040832.
PR 27-AUG-1999; 99DE-01040833.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041395.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042078.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
XX

PA (BADI) BASF AG.

Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G, Pi XY

DR WPI; 2001-071486/08.
DR N-PSDB; AAF67934.

DR N-PSDB; AAF67934.

PT Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying *C. glutamicum* or related bacteria, and as markers for
PT transformation.

PS Claim 20; Page 720-721; 1119pp; English.

The MCT nucleic acids and proteins are useful in the identification of microorganisms which can be used to produce fine chemicals, for modulating fine chemical production in C. glutamicum or related bacteria (e.g. Brevibacterium lactofermentum), the typing or identification of C. glutamicum or related bacteria, as reference points for mapping C. glutamicum genome, and as markers for transformation.

AAB68082 and AAB68082 represent sequencing primers which are used in an example from the present invention

SQ Sequence 275 AA;

Query Match	9.08;	Score 219.5;	DB 4;	Length 275;
-------------	-------	--------------	-------	-------------

best local similarity 23.1%; freq. NO. 1.9e-11;
Matches 66; Conservative 52; Mismatches 127; Indels 41; Gaps 5;

42 LKGFIFIKLMDAPLAVAAEKGFFEBDEGLFVQLAQNANKVMDRVVANGELDGSHTMLAPAR 101
 26 LTIGYVPIAGSAPIAIADALGLFKHGCVNTLKKYSGWSDLTATAYAEQLDVAHMLSPWT 85

db 26 LTIGYVPIAGSAPIAIADALGLFKHGVTTLKKYSGWSDLTWTAYATEQLDVAHMLSPMT 85

QY 102 LAAASVGF-GIKADIEVPFSGMGNGALTVSNRIMHQNKPIPLEGGKRVHPIKADYLEPV 160

Db 86 VALNAGVTNASRPTELSFTONTNGAITLASKHYGSVNSADLKG----- 130

Db 86 VAINAGVTNASRPTELTSFTQNTNGQAITLASKHYGSVNSAADLKG----- 130

101 VENTALEGKPFNMAMIFFAUSHNLIKRWLWLAGGIRNGYISPPQDISGQIGNALTSVIP 220
 131 ----MWLGIPREYSV-----HALLIRDYLVSNAPD-----IADLELRRLR 167

131 -----MVLGIPEKYSV-----HALLLRDYLVSNADP-----IADLELRLLR 167

Db 168 PADMAQLTVEGIDGFIGPFPFNEKAISNGSGRIWLLTKQLMDKHPCCAVAMAKEMKAHEH 2277

168 PADWVAQLIVEGIDGFIGPFENERAISNGSGRIWLITKQLWDXHPCCAVAMAKEWKAEN 227

Db 228 PTAAGVLNATBEASAIL--SNPAQFSSARTLSQEKYLNPATIL 271

220 FIVE V LIVES L E A S A I L - - S N E H U F D S S A K I L S Q E N I L N U P A I L L 2 / 1

Search completed: May 18, 2005, 17:42:07
Job time : 176 secs

Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 17:32:50 ; Search time 27 Seconds
(without alignments)
1282.858 Million cell updates/sec

Title: US-10-689-200-2

Perfect score: 243
Sequence: 1 MKTIRSSSKLLTLASL.....AKFAIGKQTVAGKVD 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	577.5	23.6	456	4	US-09-489-039A-10363
2	446.5	18.3	474	4	US-09-540-236-2892
3	348.5	13.9	336	4	US-09-328-352-7572
4	292.5	12.0	249	4	US-09-252-991A-30501
5	240	9.8	419	4	US-09-252-991A-25707
6	219.5	9.0	275	4	US-09-602-787A-384
7	157	6.4	197	4	US-09-252-991A-30392
8	121.5	5.0	336	4	US-09-502-540-12264
9	116.5	4.8	337	4	US-09-252-991A-26757
10	116.5	4.8	633	4	US-09-252-991A-25810
11	103.5	4.2	327	4	US-09-489-039A-12136
12	102	4.2	369	4	US-09-107-532A-5754
13	102	4.2	1062	4	US-09-397-550-4
14	102	4.2	1076	4	US-09-470-443-6
15	102	4.2	1082	4	US-09-397-550-5
16	102	4.2	1109	4	US-09-397-550-6
17	102	4.2	1115	4	US-09-397-550-23
18	102	4.2	1145	4	US-09-470-443-2
19	102	4.2	1145	4	US-09-470-443-4
20	102	4.2	1145	4	US-09-397-550-20
21	101.5	4.2	952	4	US-09-328-352-5611
22	101.5	4.2	1252	4	US-10-012-762-20
23	101.5	4.2	1252	4	US-09-704-036B-20
24	101	4.1	858	4	US-09-255-829-22
25	101	4.1	858	4	US-09-255-829-29
26	101	4.1	1169	4	US-09-255-829-20
27	98.5	4.0	315	4	US-09-248-796A-16091

28	98.5	4.0	1864	2	US-08-804-227C-3	Sequence 3, Appli
29	98	4.0	324	4	US-08-956-171E-5209	Sequence 5209, Ap
30	98	4.0	324	4	US-08-781-986A-5209	Sequence 5209, Ap
31	96.5	4.0	448	4	US-09-198-452A-216	Sequence 216, App
32	96.5	4.0	457	4	US-09-438-185A-199	Sequence 199, App
33	96.5	4.0	659	1	US-08-258-639A-4	Sequence 4, Appli
34	96.5	4.0	659	2	US-08-300-951A-4	Sequence 4, Appli
35	96.5	4.0	659	5	PCT-US95-07391A-4	Sequence 4, Appli
36	95.5	3.9	893	4	US-09-489-039A-14127	Sequence 14127, A
37	93.5	3.8	540	4	US-09-302-540-14003	Sequence 14003, A
38	93.5	3.8	1018	1	US-08-452-052-2	Sequence 2, Appli
39	92	3.8	620	3	US-09-442-100-11	Sequence 11, Appli
40	92	3.8	620	4	US-08-939-106-11	Sequence 11, Appli
41	92	3.8	620	4	US-09-442-102-11	Sequence 11, Appli
42	91	3.7	379	4	US-09-071-035-6	Sequence 6, Appli
43	91	3.7	402	4	US-09-489-039A-11633	Sequence 11633, A
44	91	3.7	431	4	US-09-134-000C-6307	Sequence 6307, Ap
45	91	3.7	777	4	US-09-917-254-57	Sequence 57, Appli

ALIGNMENTS

RESULT 1	US-09-489-039A-10363	Sequence 10363, Application US/09489039A
;	Patent No. 6610836	
;	GENERAL INFORMATION:	
;	APPLICANT: Gary Breton et al	
;	TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA	
;	FILE REFERENCE: 2709.2004001	
;	CURRENT APPLICATION NUMBER: US/09/489,039A	
;	CURRENT FILING DATE: 2000-01-27	
;	PRIOR APPLICATION NUMBER: US 60/117,747	
;	PRIOR FILING DATE: 1999-01-29	
;	NUMBER OF SEQ ID NOS: 14342	
;	SEQ ID NO 10363	
;	LENGTH: 456	
;	TYPE: PRT	
;	ORGANISM: Klebsiella pneumoniae	
;	US-09-489-039A-10363	
Query Match	23.6%; Score 577.5; DB 4; Length 456;	
Best local Similarity	33.5%; Pred. No. 2,5e-49;	
Matches 153; Conservative 76; Mismatches 173; Indels 55; Gaps 17;		
QY	SSSKKLLTLTSLASLAVGLTIAPDV-----GAVGKLEKEDLKFGFKLTDMAPLVAALAE 60	
DB	44 SISRRRLTGALG--GAMTLPGVQAAMAGSDRPEQTRVAVGFPLTDCAPLAIATA 101	
QY	61 KGFEEBGLFVQLEAANKVVDVNGELDSHMLAPPLAASVFGTKADIEVP-FS 119	
DB	102 KGFDOXYGTTLVASKASMAVADKLVAAGELDAHLTYGLYGLIEIGIAKPOAMANLMT 161	
QY	120 MGENGAITVSNINIMQKPNIPLEGKPVHPKADYLKPVKRYKAEKPFMMATFPA 179	
DB	162 LNNNGAITLSSLT--QEKVYDPLG-----LKLIDR-SAPGS-YTFATHTFT 206	
QY	180 GSHNKLRLVLAAGINPGYSPDODISQIGADALLSVTPPQMBSTLEAGTIFGCVG 229	
DB	207 GTAAWMLYVLAAGIDP-----FNDVTVVVPVPPQWNNRIGNWSGFCVG 253	
QY	240 EPNNAQAVKGVGPVITDELMKDPKVFQYTKQMAEKVPMTYLAVTALRAAIWLD 259	
DB	254 EPNNAALINDRIGFTATSDIPEPEKVLGTRRUVENNPATRLVAALMEADQRTI 312	
QY	300 ADNNKRRKEALEWIAKQYVADGVLAASMGNTFEYKODKRALPDFN--TFRRHG-AS 356	
DB	313 AASPENTRETRARLARARCMINTKEQVLTGRLG--EYDNGLGRRMKDAHIRRWAGGEVS 370	
QY	357 YPSYSAVWYLTQLRWGMINEFKPDNNWYLDIAKNVYRPDIYLAALKEVLAAEKAAEDF 416	

Db 371 FFWLSDGMWFLTOFRBMLKQ-APD--YLAVASRINRIDWQAAQ---AVGISA--- 421
 QY 417 PADTSLKPSQNFIDKVPDPANKPNYLAKFAIGLG 453
 Db 422 PA---ARMRSSTLMGDTVMWNGSDPEGARHFSIORKG 455

RESULT 2

US-09-540-236-2892
 ; Sequence 2892, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540,236
 ; CURRENT FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 2892
 ; LENGTH: 474
 ; TYPE: PRT
 ; ORGANISM: M. catarrhalis
 ; US-09-540-236-2892

Query Match 18.3%; Score 446.5; DB 4; Length 474;
 Best Local Similarity 27.9%; Pred. No. 4, se-36;
 Matches 159; Conservative 82; Mismatches 180; Indels 71; Gaps 11;

QY 2 KTIIRSSSKLLTLTSSASLAWGLTAPDVGAGKL--EKEDLKPFGIKLTDMAPLAVA 59
 Db 79 KTVGKTTA-----LAIASVLPALTLQEAADTLKPKKSDVIGFLPLCTPLIMAD 132
 QY 60 EKGFEDEGLFVQLAQAAMKVVMDRVNGLDGHMLAPAPLASVGFGT-KADIEVP 118
 Db 133 PLGYVAEGIKANLTKRAGMALVRDQMRBELDAHPLAMPAINMIGLSAKQNKVAA 192
 QY 119 SMGFENGLATTVSNEIWHQKKNPRLGSGKRVHRIKADYIKRVVEKTKAGKPPNMAITP 178
 Db 193 IONTNGQALVMA-----LKHNNRNPKNWKGTFTALPFE 226
 QY 179 AGSHNKLKRWLAAGINPGYSPPODISGADALSLVTPPPQMPSTLEAGTIGYCV 238
 Db 227 HSHHNYLAFYLAHSLDP-----DKDKALITTPPMIAMLAKGNIDGFFG 273
 QY 239 GEPNQAQVFKGIGVITDEELMKOTPEKVGVTQMAKTPNTYLAVTAKALIRAIWL 298
 Db 274 PEPFQRAVWDKAGYIHTLSRIDMNGHPCCSFQTSQSFINDYQETPLAMYRAIKANVM- 332
 QY 299 DADNNKREALEMLAQKYGADVLAASMGTFEYKDKKALPDNTPFRHGAS-Y 357
 Db 333 -ANKDSIRKDLKLSPAYVLAQPELVLSQSIWGRFADGVDTIQVDP-----RMGFDM 386
 QY 358 PSYSSAVWYLTQLRKGMINEFKPDWYLDTAKNVPRPIYLAALAEKAKADEPP 417
 Db 387 PMSHVAAMWMTQMKRGYIT--GNINYODIANQV--MLDAKQOMQAMGYTVVQDEP 440
 QY 418 ADTSIKPSQNFIDKVPDPANKPNYLAKFAIGLKQTFVAG 459
 Db 441 KITVMGKQ-----FNATDPDAYLDSFAIGHKSTGRLHG 473

RESULT 3
 US-09-328-352-7572
 ; Sequence 7572, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 7572
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-7572

Query Match 13.9%; Score 338.5; DB 4; Length 336;
 Best Local Similarity 26.5%; Pred. No. 2e-25;
 Matches 96; Conservative 63; Mismatches 144; Indels 59; Gaps 9;

QY 34 VGKLEKEDLKFGFIKLTDMAPLAVAEEKPFDEGLFVQLBAQAMKVVMDRVNGLDQ 93
 Db 1 MSKLETTQQLGTYPLDLCIALMAQGFEEVLDVTLVKEASASLRDLAFELDA 60
 QY 94 SHMLAPAPLASVGFCTKAD-----IEVPSMGFNATVSNELWHQW---KPNIPLE 144
 Db 61 AHCLSMLPRAAMG-----ADQIGIALQTLVLSKRNAPISLSQKLIHQALIKENDVAQTT 116
 QY 145 GSKPVHPIKADYIKRVVEKTKAGKPPNMAITPFRHGASYPSSAAWYLTQLRKGMINEK 204
 Db 117 AOKVIOYIEQDH-----TLSLAHVFKSHHHCLEWMLA----- 152
 QY 205 DISGQIGADALSLVTPPPQMPSTLEAGTIFGYCVGEPNQAQVFKGIGVITDEELMKD 264
 Db 153 --DSRIAQTLKALKALPFPYMWBALDNHVIDGFCVGEPPMTQGBLGLSKIVCSQDIIFN 210
 QY 265 TPEKVGVTQMAKTPNTYLAVTAKALIRAIWLDADNNKREALEMLAQ---KQYVG 320
 Db 211 VADKVLAVTQEWAEQHPTLVALTAIMKAQ--KELSNLKDPAPVYKLVFEGIVFHC 268
 QY 321 ADVEYLAASMGTFEYKDKKALPDNTPFRHGASYPSSAAWYLTQLRKGMINEK 380
 Db 269 EEVHNDKRYMIQNIYKHLVENAAPPQEDFH-----MLFQOMQWKEKL-QLA 314
 QY 381 PD 382
 Db 315 PD 316

RESULT 4
 US-09-252-991A-30501
 ; Sequence 30501, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30501
 ; LENGTH: 249
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-30501

Query Match 12.0%; Score 292.5; DB 4; Length 249;
 Best Local Similarity 31.3%; Pred. No. 5, 4e-21;
 Matches 73; Conservative 40; Mismatches 105; Indels 15; Gaps 6;

QY 218 VTPPQMPSTLEAGTIFGYCVGEPNQAQVFKGIGVITDEELMKOTPEKFGVTQMA 277
 Db 19 VVPQWQGHLOAGRIDGCGAPWGAALVDDGQGTATATSOAIWPDHPEKVLGTTAFV 78
 QY 278 EKPNTYLAVTAKALIRAIWLDADNNKREALEMLAQKYGADVVAASMGTFEY 337
 Db 79 DAYPNTAALVMAVLDASRFE-QVAENRLGTRQLISGRDYDAPALGAIQPFPGYQDG 137

APPLICANT: Zelder, Oskar
 APPLICANT: Habermann, Gregor
 TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
 TITLE OF INVENTION: TRANSPORT
 FILE REFERENCE: BGI-125CP
 CURRENT APPLICATION NUMBER: US/09/602,787A
 CURRENT FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US/N 60/141031
 PRIOR FILING DATE: 1999-06-25
 PRIOR APPLICATION NUMBER: DE 199331454.3
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 199331478.0
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 199331563.9
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 199332122.1
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332124.8
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332125.6
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332128.0
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332180.9
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332182.5
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332190.6
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 199332191.4

```

1 PRIOR APPLICATION NUMBER: DE 199332209.0
2 PRIOR FILING DATE: 1999-07-09
3 PRIOR APPLICATION NUMBER: DE 199332212.0
4 PRIOR FILING DATE: 1999-07-09
5 PRIOR APPLICATION NUMBER: DE 199332227.9
6 PRIOR FILING DATE: 1999-07-09
7 PRIOR APPLICATION NUMBER: DE 199332228.7
8 PRIOR FILING DATE: 1999-07-09
9 PRIOR APPLICATION NUMBER: DE 199332229.5
10 PRIOR FILING DATE: 1999-07-09
11 PRIOR APPLICATION NUMBER: DE 199332230.9
12 PRIOR FILING DATE: 1999-07-09
13 PRIOR APPLICATION NUMBER: DE 199332227.3
14 PRIOR FILING DATE: 1999-07-14
15 PRIOR APPLICATION NUMBER: DE 199333005.0
16 PRIOR FILING DATE: 1999-07-14
17 PRIOR APPLICATION NUMBER: DE 199333006.9
18 PRIOR FILING DATE: 1999-07-14
19 PRIOR APPLICATION NUMBER: DE 19940764.9
20 PRIOR FILING DATE: 1999-08-27
21 PRIOR APPLICATION NUMBER: DE 19940765.7
22 PRIOR FILING DATE: 1999-08-27
23 PRIOR APPLICATION NUMBER: DE 19940766.5
24 PRIOR FILING DATE: 1999-08-27
25 PRIOR APPLICATION NUMBER: DE 19940830.0
26 PRIOR FILING DATE: 1999-08-27
27 PRIOR APPLICATION NUMBER: DE 19940831.9
28 PRIOR FILING DATE: 1999-08-27
29 PRIOR APPLICATION NUMBER: DE 19940832.7
30 PRIOR FILING DATE: 1999-08-27
31 PRIOR APPLICATION NUMBER: DE 19940833.5

```

PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941395.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942077.7
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942078.5

ORGANISM: Homo sapiens
US-09-470-443-6

Query Match
Best Local Similarity 23.7%; Score 102; DB 4; Length 1076;
Matches 79; Conservative 42; Mismatches 142; Indels 70; Gaps 17;

QY 7 SSSKLLLTLSASLAWGLTIAPDVAGVCKLEKEDLKFGFIKLTDMAPLAVAEKGFED 66
DB 218 SSPKDMVITVDVSGSVGLTL-----KLMKTSVCEMLDTLSDDDYVNVASFNEKAQP 269
QY 67 EGI.FVOL-BAQANWKYVMDRVNNGELDGSIMLAPPLAASVGFRTKADIEVPPSMGFNGN 125
DB 270 VSCFTHLVQANVANKKVFKEAVG-----MVAKGTGYKAGFEYAFDQLQNSN 317
QY 126 AITVSEIHWOMKPNIP.L--EGSKPVHPIKADYLKPYVEKYKABGKPFNNAMTFPAGSHN 183
DB 318 -ITRAN-----CNKMTIMPTDGE-----DRVQDFEKTNNPNNRTVRV-FTFSVGQHN 363
QY 184 IKLR--YMLAAGINPGYSPPODISG-OIGADALLSVTPPQMPSTLEAGTIFGYCVGE 240
DB 364 YDVTPLQWMAA--NKGYPFIRPSIGAIRINTQEVLDVGRPMVLAGKEAKQV----- 414
QY 241 PMNQAVFK-GIGVPVITDEBELMKDTPBEKVPFVTKQMAEKYPNTYLAVTKALIRAAIWL 299
DB 415 QWNTN--VYEDALGLGLVVTGTL-----PVFNLTODGPEKKN-----QLILGWGID 459
QY 300 ADNNKRRKEAIEMLAQKQYVADVEYLAASMNG 332
DB 460 VALND-----IKRLTPNTYLGANGYVFAIDING 487

RESULT 15
US-09-397-550-5
; Sequence 5, Application US/09397550
; Patent No. 6783952
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert
; TITLE OF INVENTION: Secreted soluble alpha2 delta-2, alpha2 delta-3, alpha2
; TITLE OF INVENTION: delta-4 calcium channel subunit polypeptides and
; FILE REFERENCE: 180
; CURRENT APPLICATION NUMBER: US/09/397,550
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1082
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-397-550-5

Query Match
Best Local Similarity 23.7%; Score 102; DB 4; Length 1082;
Matches 79; Conservative 42; Mismatches 142; Indels 70; Gaps 17;

QY 7 SSSKLLLTLSASLAWGLTIAPDVAGVCKLEKEDLKFGFIKLTDMAPLAVAEKGFED 66
DB 287 SSPKDMVITVDVSGSVGLTL-----KLMKTSVCEMLDTLSDDDYVNVASFNEKAQP 338
QY 67 EGI.FVOL-BAQANWKYVMDRVNNGELDGSIMLAPPLAASVGFRTKADIEVPPSMGFNGN 125
DB 339 VSCFTHLVQANVANKKVFKEAVG-----MVAKGTGYKAGFEYAFDQLQNSN 386
QY 126 AITVSEIHWOMKPNIP.L--EGSKPVHPIKADYLKPYVEKYKABGKPFNNAMTFPAGSHN 183
DB 387 -ITRAN-----CNKMTIMPTDGE-----DRVQDFEKTNNPNNRTVRV-FTFSVGQHN 432
QY 184 IKLR--YMLAAGINPGYSPPODISG-OIGADALLSVTPPQMPSTLEAGTIFGYCVGE 240
DB 433 YDVTPLQWMAA--NKGYPFIRPSIGAIRINTQEVLDVGRPMVLAGKEAKQV----- 483
QY 241 PMNQAVFK-GIGVPVITDEBELMKDTPBEKVPFVTKQMAEKYPNTYLAVTKALIRAAIWL 299

DB 484 QWNTN--VYEDALGLGLVVTGTL-----PVFNLTODGPEKKN-----QLILGWGID 528
QY 300 ADNNKRRKEAIEMLAQKQYVADVEYLAASMNG 332
DB 529 VALND-----IKRLTPNTYLGANGYVFAIDING 556
Search completed: May 18, 2005, 17:42:39
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 17:39:18 ; Search time 144 Seconds
(without alignments)
1077.858 Million cell updates/sec

Title: US-10-689-200-2
Perfect score: 2443
Sequence: 1 MKTIIRSSSKLLTLTSLASL.....AKFAIGLKQGVAVGAKVD 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seags, 334507595 residues
Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: Published Applications AA.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2443	100.0	464	US-10-689-200-2	Sequence 2, Appl1
2	525	21.5	403	US-10-282-122A-69975	Sequence 69975, A
3	240	9.8	399	US-10-282-122A-66370	Sequence 66370, A
4	240	9.8	399	US-10-389-647-502	Sequence 502, App
5	234.5	9.6	396	US-10-282-122A-49807	Sequence 49807, A
6	227	9.3	396	US-10-282-122A-44527	Sequence 44527, A
7	226	9.3	394	US-09-738-626-4906	Sequence 4906, Ap
8	222.5	9.1	372	US-10-282-122A-68363	Sequence 68363, A
9	220.5	9.0	391	US-10-282-122A-49957	Sequence 49957, A
10	219.5	9.0	375	US-10-627-476-384	Sequence 384, App
11	217.5	8.9	391	US-10-282-122A-48182	Sequence 48182, A
12	144	5.9	342	US-10-282-122A-72603	Sequence 72603, A
13	140	5.7	192	US-10-282-122A-47676	Sequence 47676, A

14	116.5	4.8	883	14	US-10-156-761-8511	Sequence 8511, Ap
15	111.5	4.6	333	14	US-10-156-761-9969	Sequence 9969, Ap
16	108	4.4	533	15	US-10-425-114-51223	Sequence 51223, A
17	107	4.4	518	15	US-10-424-599-198653	Sequence 198653, A
18	106.5	4.4	461	15	US-10-282-122A-46659	Sequence 46659, A
19	105	4.3	419	15	US-10-282-122A-46437	Sequence 46437, A
20	105	4.3	1130	15	US-10-275-595A-6	Sequence 6, Appl1
21	105	4.3	3352	14	US-10-156-761-7961	Sequence 7961, Ap
22	105	4.3	4455	15	US-10-287-226-304	Sequence 304, App
23	104	4.3	866	14	US-10-241-596-104	Sequence 104, App
24	104	4.3	1420	14	US-10-241-596-110	Sequence 110, App
25	103.5	4.2	280	9	US-09-815-242-11714	Sequence 11714, A
26	103	4.2	864	14	US-10-241-596-102	Sequence 102, App
27	102.5	4.2	320	15	US-10-282-122A-55774	Sequence 55774, A
28	102	4.2	613	15	US-10-369-493-18239	Sequence 18239, A
29	102	4.2	860	14	US-10-241-596-175	Sequence 175, App
30	102	4.2	862	14	US-10-241-596-94	Sequence 94, App
31	102	4.2	866	14	US-10-241-596-88	Sequence 88, Appl
32	102	4.2	867	14	US-10-241-596-96	Sequence 96, Appl
33	102	4.2	867	14	US-10-241-596-98	Sequence 98, Appl
34	102	4.2	870	14	US-10-241-596-92	Sequence 92, Appl
35	102	4.2	871	14	US-10-241-596-84	Sequence 84, Appl
36	102	4.2	871	14	US-10-241-596-86	Sequence 86, Appl
37	102	4.2	871	14	US-10-241-596-90	Sequence 90, Appl
38	102	4.2	1062	17	US-10-902-531-4	Sequence 4, Appl1
39	102	4.2	1076	14	US-10-116-949-6	Sequence 6, Appl1
40	102	4.2	1082	17	US-10-902-531-5	Sequence 5, Appl1
41	102	4.2	1109	17	US-10-902-531-6	Sequence 6, Appl1
42	102	4.2	1115	17	US-10-902-531-23	Sequence 23, Appl
43	102	4.2	1145	14	US-10-116-949-2	Sequence 2, Appl1
44	102	4.2	1145	14	US-10-116-949-4	Sequence 4, Appl1
45	102	4.2	1145	17	US-10-482-029-114	Sequence 114, App

ALIGNMENTS

RESULT 1
US-10-689-200-2
; Sequence 2, Application US/10689200
; Publication NO. US20040126848A1
; GENERAL INFORMATION:
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Ni, Hao
; APPLICANT: Ye, Rick
; APPLICANT: picataggio, Stephen
; APPLICANT: Wang, Iao
; APPLICANT: Seip, John E.
; TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN CI METABOLIZING BACTERIA
; FILE REFERENCE: CL1747 US NA
; CURRENT APPLICATION NUMBER: US/10/689,200
; PRIOR FILING DATE: 2003-10-20
; PRIOR FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Methylobionas sp. 16a
; US-10-689-200-2

Query Match 100.0%; Score 2443; DB 16; Length 464;
Best Local Similarity 100.0%; Pred No. 2.5e-216;
Matches 464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIIRSSSKLLTLTSLASLAVGLTIADVGVAGKLEKEDLKFGRFKLTDMAPLAAAE 60
DB 1 MKTIIRSSSKLLTLTSLASLAVGLTIADVGVAGKLEKEDLKFGRFKLTDMAPLAAAE 60
QY 61 KGFPEDEGLFVQLEADANKVMVDRVNVGELDGSMLAPAPLAAVGFGRKADIEVPPFSM 120
DB 61 KGFPEDEGLFVQLEADANKVMVDRVNVGELDGSMLAPAPLAAVGFGRKADIEVPPFSM 120

```

QY 121 GENGNAITVSNELTHQMKPNIPLEGKPVHPRIKADYLKPVVEKXKAEKGFNNMAMTFPAG 180
DB 121 GENGNAITVSNELTHQMKPNIPLEGKPVHPRIKADYLKPVVEKXKAEKGFNNMAMTFPAG 180
QY 181 SHNIKRLWYLAAGGIPGYSPPODISGOIGADALSTVTPPPMPSTLEAGTIFGCVGB 240
DB 181 SHNIKRLWYLAAGGIPGYSPPODISGOIGADALSTVTPPPMPSTLEAGTIFGCVGB 240
QY 241 PNNQAVFVGKIGVPIVTTDELMKOTPEKXVGVTKQMAEKYPNTYLAVTKALIRAAIWLDA 300
DB 241 PNNQAVFVGKIGVPIVTTDELMKOTPEKXVGVTKQMAEKYPNTYLAVTKALIRAAIWLDA 300
QY 301 DNNKRRKALIMLAQOYVIGADVETLAASNGTFEEFKODKALPDPNTFFRGGASYPST 360
DB 301 DNNKRRKALIMLAQOYVIGADVETLAASNGTFEEFKODKALPDPNTFFRGGASYPST 360
QY 361 SSAVWYLTQRRKGMINEKPDNMYLDTAKNVPYRPIYLAARELAEGKAKAEDPADT 420
DB 361 SSAVWYLTQRRKGMINEKPDNMYLDTAKNVPYRPIYLAARELAEGKAKAEDPADT 420
QY 421 STKPSQNFIDKVPDPANKENDYLAFAIGLKGKQTVAGSKVVD 464
DB 421 STKPSQNFIDKVPDPANKENDYLAFAIGLKGKQTVAGSKVVD 464

```

RESULT 2

```

US-10-282-122A-69975
/ Sequence 69975, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:

```

```

/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Foreyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 69975
/ LENGTH: 403
/ TYPE: PRT

```

```

/ ORGANISM: Pseudomonas syringae
/ US-10-282-122A-69975

```

```

Query Match 21.5%; Score 525; DB 15; Length 403;
Best Local Similarity 31.7%; Pred. No. 26-39;
Matches 132; Conservative 72; Mismatches 164; Indels 48; Gaps 12;

```

```

QY 38 EKEEDLKFGFTIKLTDMAPLVAAREKGFEEDEGIFVQLEADANKVMDRVYNGELDGSNML 97
DB 19 ENMSLDVGFPMALTDCAPLVVAATOGFAOPGYSLMKROTSSWAGLDRILVSQLOQAHSI 78
QY 98 AAPLAASVGF--GTRKADIEVPFSGMGNALTVSNELTHQMKPNIPLEGKPVHPRIKAD 155
DB 79 YGLIYAVEIGISGPRPTDAIIMGLNQGCTINLSRE-----LDDAGVITPEALD 128
QY 156 YLKPVEKXKAEKGFNNMAMTFPAGSHNIKRLWYLAAGGIPGYSPPODISGOIGADAL 215
DB 129 -----KRAHGGSKLTFKQTFPTGNHAWMLYYMLASQIHP-----LDDV 168
QY 216 LS-VTPPPMPSTLEAGTIFGCVGBEPNNQAVFVGKIGVPIVTTDELMKOTPEKXVTK 274
DB 169 TSVVVPPTQMAOHLAAGRIDGFCVGBPWGASAVQOOLGFTWATSOAIWPDHKGVLGCTR 228
QY 275 QMAEKYPNTYLAVTKALIRAAIWLDA DNNKRRKALIMLAQOYVIGADVETLAASNGTF 334
DB 229 EFEQNPNTARALIMAVLASRFIE-QSDHNRSTRQULSGVDYLDASLDCTIEPRLLGY 287
QY 335 EYKODKRALPDNTFFRGG-ASYPSSAVWYLTQRRKGMINEKPDNMYLDTAKNYY 393
DB 288 SDGLGNQWMDPRAVSHFDQGVNYFPLSDGMFMFMQFRWGLRE-DPD-YLAVASRVQ 344
QY 394 RPDYLAARELAEGKAKAEDFPADTSTKPSQNFIDKVPDPANKENDYLAFAI 449
DB 345 QLDLYROANAL-----GIDAPSAT-LRSSQ--LIDGKVDGSDPAGYARSFKL 390

```

RESULT 3

```

US-10-282-122A-66370
/ Sequence 66370, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:

```

```

/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Foreyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22

```



```

; SOFTWARE: Patentin version 3.1
; SEQ ID NO 49807
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
; US-10-282-122A-49807

Query Match
Best Local Similarity 23.3%; Score 234.5; DB 15; Length 396;
Matches 105; Conservative 77; Mismatches 150; Indels 119; Gaps 18;

QY 9 SKKLLTLLTSASLAWGLTTPADVGAAGKLEKED---LKEGFIKLTDMAPLAAAKGFF 64
DB 7 SREEMKXASMTFVAG--AAPLLAALNAAENRDPAPRIGYPTIDAPLVAHNNGYF 64
QY 65 EDEGLFVQLA--QANKVMDRVVNGELDGSMLADAPLAASVFGTKADIEVFGMGFN 123
DB 65 DASGLAVEKPTLLRSWAQVLEAFISQVNVVHLLAMTWTAR--YSGQAPAKVAMNHYN 122
QY 124 GNAITVSNITWQMKNNIPLEGKRPVPIKADYIKVVEKYKAEGRPMAMTFPAGSHN 183
DB 123 GSALTVAPI-----GKLELGGKTV-----AVPEWYSIHN 153
QY 184 IKLRWYLAAGINPGYSPQDISQIGADAL-LSVTPPQMPSTLEAGTIFGVCGBPM 242
DB 154 VVVGHLRAQGL-----VPLEKQGLKANEVRLIMSSDMPALASQIAGFIYAEFF 208
QY 243 NOQAVFKGIGVPTIDEELMKD-----TPKVFQVTKQMAEKYDNTYLAATKALIRAA 295
DB 209 NAAAEELKVGKYLRFQGDVWKNHACCVEFMHERDLTERAAMSQK-----VVDVAVKIQ 261
QY 296 IMLDAMNNKREKATEMLA---OKQVGADEVYLAASNGTFEYKDDKRALPDFTFFR 352
DB 262 VWTRA---HQBAAQLLSKSGNHYTPHSANVLT---VLAAPPGDEBRVYAD----- 308
QY 353 HGASYPSSAAVWYLTQLRWGMINEFKPDNWTLDYAKV--YRPDIYLAARELV----- 406
DB 309 -----RAIHHADWH-----AKRIDQPYRYPAYTEBELVRLKA 341
QY 407 --AEGKAK-----ADEPPADTSIRPS 425
DB 342 TQVEGNAQFLHQLDPAFVARDLVDDRFFVKS 372

RESULT 6
US-10-282-122A-44527
; Sequence 44527, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
```

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 44527
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; FEATURE: MISC FEATURE
; NAME/KEY: (7)..(7)
; LOCATION: (7)..(7)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-44527

Query Match
Best Local Similarity 9.3%; Score 227; DB 15; Length 257;
Matches 69; Conservative 52; Mismatches 122; Indels 50; Gaps 7;

QY 91 LQSHMLAPAPLAASVFGTKAD-----IEVPSMGFNATVSNITWQMKNNIP 141
DB 1 LBAAGXSMPLPAAAMG-----ADQIGIALQTPVLTSKRAFLISOKLYQLAIGESDNA 56
QY 142 PLEGKRPVPIKADYIKVVEKYKAEGRPMAMTFPAGSHNIKLRWYLAAGINPGYYS 201
DB 57 QTTAQKVIQYIQDH-----TSLAHVFKSHIHVCLREWLALA----- 95
QY 202 PPQDISQIGADALLSVTPPQMPSTLEAGTIFGVCGBPMQOAVFKGIGVPTIDEEL 261
DB 96 -----DSRLAQTLKIKALPPPYWALDNHVIDGFCVGBPMWQBELGLSKIVCSSODI 150
QY 262 WKDTEKVGTVKQMAEKYDNTYLAATKALIRAAIWLMDAMNNKREKATEMLAQKQYGA 321
DB 151 IPNVADKVLAVTQENABEQHPQTLVALLTRIMAQ--KEISNLKDPAPILKLT----- 200
QY 322 DVEVLAASNGTFEYKDDKRALPDFTFFRHGASYPSSAAVWYLTQLRWNG 374
DB 201 -VEFGIVRHCSEVHVDRKYWIMQNIYKYLKVENAPQEDPFHMLFQGWQKWG 252

RESULT 7
US-09-738-626-4906
; Sequence 4906, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASAYO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
```


PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4906
LENGTH: 294
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4906

Query Match 9.3%; Score 226; DB 9; Length 294;
Best Local Similarity 23.6%; Pred. No. 4.8e-12; Indels 44; Gaps 6;
Matches 78; Conservative 56; Mismatches 152;

QY 1 MKTIIRSSSKLLTLTSLASLAWGLTAPDVAGVGLKEKD--LKEGFIKLTMDAPLAV 57
DB 1 MTHILFDSRFLQIGAFASLSTLAGARVYISTNNRPDNPPLTIGVPIASAPIAI 60
QY 58 AAEKGFEEDEGLFVQLEAOANMKVMDRVVNGELDSHMLAPPLASVGF-GTKADIEV 116
DB 61 ADALGIFKGGVAVTLTKYSGWMDLWTAYATEQLDVAHMLSPMTVAINGVTNASRPTL 120
QY 117 PSMGNGNAITYSNIMHQMKNIPLEGKPVHPKADYLKRVKRYAEKGFPMNMT 176
DB 121 SFTQNTGQATITLASKHYSSVNSADLKG-----MVLGIPFYSV- 160
QY 177 PFGSHNIKRLRYWLAAGINPGYSPPODISQIGADALISVTPPPMPSTLEAGTIFGY 236
DB 161 -----HALLRDLYVSNANP-----TADLELRIRPDMAQLTVEGIDGF 202
QY 237 CVGEPNQAVFKGIGVPVITDELMKDPPEKVFVGTOKMAEKYPTTYLAVTALIRAI 296
DB 203 ICPGPENERAISNGSGRIWLTQKQMDKPCCAVMAKEMKAEHPTAAGVINALBEASA 262
QY 297 WLDADNNKREKRIEMLAQKQYGADEVYL 326
DB 263 IL--SNPAQFDSSARTLSQEKYILNQPATLL 290

RESULT 8

US-10-282-122A-68363
Sequence 68363, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 68363
LENGTH: 372
TYPE: PRT
ORGANISM: Pseudomonas putida
US-10-282-122A-68363

Query Match 9.1%; Score 222.5; DB 15; Length 372;
Best Local Similarity 23.1%; Pred. No. 1.4e-11; Indels 89; Gaps 17;
Matches 90; Conservative 71; Mismatches 139;

QY 42 LKEGFIKLTMDAPLAVAAEKGFEEDEGLFVQLEA-QANMKVMDRVVNGELDSHMLAPA 100
DB 17 VRIGVLPITDATTLLVAHNNGLFEABGIRKAEVILRSMAQVIEAFISQVNVYHLLSPM 76
QY 101 PLAASVGFETKADIEVPFSKGFNGNAITYSNEIMHQMKNIPLEGKPVHPKADYLKPV 160
DB 77 TWVAR--YGSKVPAKVANNHVGSGITVAPDI-----SAVKQIGGKTV----- 118
QY 161 VEKRYAEKGFPMNMTFPAGSHNIKRLRYWLAAGINPGYSPPODISQIGADAL-LSVT 219
DB 119 -----ALFHWYSIHNVVLQQLNDNGLRP--VSKP--ANAQLAANEVNLVL 161
QY 220 PPPMPSTLEAGTIFGYCVGEPNQAVFKGIGVPVITDELMKD-----TPEKVFV 272
DB 162 PPEDMPALASKRIAGTYAEFPNALAENLKVRVQRTGVDVRNACCVFFHEHDLN 221
QY 273 TKQMAEKYPTTYLAVTALIRAIWLDADNNKREKRIEMLAQ--KQYGADEVILAS 329
DB 222 RPEWSQK-----VYNAIVKAQW-----TRDRTEAALLSRAGPKYTPHPBAVLTKV 270
QY 330 M-----NGTFEYKDKRALPDNTEFRIGASIPSSAVMYLTQLRMGMIN 377
DB 271 LAPAEDRAGYIASGAIRHQWDEKRI-DFQPY-----PPSYTER---LVKRLKTLTI- 320
QY 378 EFKPDNMYLDTAKNVYRPDIYLAARELY 406
DB 321 --EGDNTFISGLDPAY-----AARDLV 340

RESULT 9

US-10-282-122A-49957
Sequence 49957, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 49957
LENGTH: 391
TYPE: PRF
ORGANISM: Burkholderia fungorum
US-10-282-122A-49957

Query Match 9.0%; Score 220.5; DB 15; Length 391;
Best Local Similarity 24.4%; Pred. No. 2,4e-11;
Matches 94; Conservative 63; Mismatches 136; Indels 93; Gaps 16;

9 SKKLLTTLTASLAWGLTAPDGAIVGKLEKED---LKEFGIKLTDMAPLVAAEKGF 64
2 SREMLKIASMTVAG--AGPLIGALNAAADDPNAPVIGIHPITDPTLVVANNNGF 59
65 EDEGLFVQLEA-QANMKVMDRVVNGELDSHMLAPAPLAASVGTCTADIEVPSMGFN 123
60 EAAGLAVEKPTLLRSVAQLVEAFVLSQVNVVHLAPMTIAR--YGSQAPAKVVMNNHN 117
124 GNAITVSNIEIMQKNPILLEGKRPVPIKADYLKVEKYAEKGFPMNMTFPAAGSN 183
118 GSALTIVSPDI-----GKVADLGKTV-----AVPFWYSIHN 148
184 IKLRFWLAAGINP-----GYSPQDISGQIGADALLSVTPPQWSTLEAGTFGCV 238
149 VVVQMLRQGLVPLVKDGEIKP-----NEKRLVMAPSDIPPLASRQJAGFTV 199
239 GEPNQAQVFKSIGVPTIDELMKD-----TEKVPGVTKQWAEKYPTNYLAATKAL 291
200 AEPFNAABGLKIGKVLRTGDMVKNHACCVFPMHERDLTERAANSQK-----VDSV 252
232 IRAATWLDADNNKRNKEAIEMLA---QKQY---VGADVEVLAASNGTFEYKDKRAL 344
253 VQAQIMTRA---HFOEAQQLSKTGQNMHTPHTAVALTEVLAPAPGQGRYIAD--RAI 306
345 -----PDNTFFRHGASVPSYS 361
307 IHADWAKRIDQPY-----PYPAIT 327

RESULT 10
US-10-627-476-384
Sequence 384, Application US/10627476
Publication No. US20040030116A1
GENERAL INFORMATION:
APPLICANT: Pompejus, Mark
APPLICANT: Kroger, Burkhard
APPLICANT: Schoder, Hartwig
APPLICANT: Zelder, Oekar
APPLICANT: Habernauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
TITLE OF INVENTION: TRANSPORT
FILE REFERENCE: BGI-125CPN

CURRENT APPLICATION NUMBER: US/10/627,476
CURRENT FILING DATE: 2003-07-25
PRIOR APPLICATION NUMBER: 09/602,787
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: USSN 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931454.3
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931478.0
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931563.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932122.1
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932124.8
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932180.9
PRIOR FILING DATE: 1999-07-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 678
SEQ ID NO 384
LENGTH: 275
TYPE: PRF
ORGANISM: Corynebacterium glutamicum
US-10-627-476-384

Query Match 9.0%; Score 219.5; DB 15; Length 275;
Best Local Similarity 23.1%; Pred. No. 1.7e-11;
Matches 66; Conservative 52; Mismatches 127; Indels 41; Gaps 5;

42 LKFGFIKLTMDPLVAAEKGFEEDEGLFVQLEAQAQNMKVMDRVVNGELDSHMLAPAP 101
26 LTIGVPIAGSAPIMADLGLFKHGVNVTLKXSGMSDMLTATATBEDLVANHLSPMT 85
102 LAASVGF-GTKADIEVPSMGFNVAITVSNIEIMQKNPILLEGKRPVPIKADYLKPV 160
86 VAINAGVTNARSPTLSFTQNTNGQAITLASKHGVSNSADLKG----- 130
161 VEKYAEKGFPMNMTFPAAGSNITLRYWLAAGINPGYSPQDISGQIGADALLSVTP 220
131 ---MWLGIPFYSV-----HALLRDTLVSNADP-----IADLEIRLIR 167
221 PQWSTLEAGTFPGYCVGEPNQAQVFKSIGVPTIDELMKDTEPEKFGVTKQWAEKY 280
168 PADWQAQLVBESIDFIFGPGFNERAIGNSGRIWLTQKQMDKRPCCAVMAKEMKAEH 227
281 PNTYLAATKALIRAAIWLADNNKRNKEAIEMLAQAQNVGADVEVL 326
228 PTAQGVNALAEBSAIL--SNPAQFDSASRTLSQEKYLNQPATLL 271

RESULT 11
US-10-282-122A-48182
Sequence 48182, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 48182
LENGTH: 391
TYPE: PRT
ORGANISM: Burkholderia cepacia
US-10-282-122A-48182

Query Match 8.9%; Score 217.5; DB 15; Length 391;
Best Local Similarity 23.3%; Pred. No. 4,5e-11;
Matches 95; Conservative 67; Mismatches 130; Indels 115; Gaps 16;
9 SKKLLTSLASLAWG---LTTA-----PDVGAVKLEKEDLKFGFIKLTMDAPLAV 57
2 SRREWLKLSALFTATGAAPLISMANAHAAEPDA-----FVRIGYLPITDAADLLV 52
58 AAKGPFEDGLTV-QLEAQAANKVVDVNGELDSHMLAPPLAASGFGKADIEV 116
53 AHNGGFASGGLTVEQPKLRISMAQVLEAPLSQGVNVHLLSPWTLMAR--YGSRAKAV 110
117 PFMGNGENALITSNEIWHQMKPNIPLEGKPVHPIKADYLKPVVEKYKAEKPFNNAMT 176
111 VAMNHVNGSGLTTPA---DSLRDLGKTV-----AVP 141
177 FPAASHNIKLRVWLAAGINPGYVSPQDISQIGADALLSVTPPQMPSTLEAGTIFGY 236
142 FWYSIHVVVLQDMLRGGLVPLKRGTPRANENVN---LVVMAFSDMLPALAARQAGY 197
237 CUEPNNQAVFVGIGVPTVTDDELMDOTBEKVFQVTKQ-----WAEKPTTYLAATK 289
198 IVAFEPFATLEMKIGILRFTGVDVKNHACCVAFMEHODLJORPWSOK-----VVN 250
290 ALIRAAIWLADNNKKEALIEMLAQOYGVADVEVLAASNGCFEYEXD-----DKRAL 344
251 AIYKAQDMA---RSHQETRAQL-----SKDGHRSPTTLASLDKVLV 291
345 PD---FNTFPRHGA-----SYSPYSAAVWYLTQLRR 372
292 PSASLADTVRASGIRHADWHAKRIDFPYRPFPSYTEAL--VQRLKR 336

RESULT 12
US-10-282-122A-72603
Sequence 72603, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 72603
LENGTH: 342
TYPE: PRT
ORGANISM: Streptococcus mutans
US-10-282-122A-72603

Query Match 5.9%; Score 144; DB 15; Length 342;
Best Local Similarity 21.4%; Pred. No. 0.00022;
Matches 83; Conservative 59; Mismatches 145; Indels 100; Gaps 15;
12 LLLTSLASLAWGTLTPADVAGVLEKEDLKFGFIKLTMD--APLAAAEKGFPEBGL 69
15 ILPLISVSLVACSKSSTSTSKNYNASEKDSYILKISNSDLGAPQOIAIEKGFDDVGL 74
70 FVQLEAQAANKV-----MDRVNNGELDSHMLAPA---PLAASVGFQTKADIEVP 118
75 -----KYKVKYLGODTSLNLDNAAGKIDASNSLMASTIOPLA----- 111
119 SMGFGNNAITVSNBIWHQMKPNIPLEGKPVHPIKADYLKPVVEKYKAEKPFNNAMT 178
112 -----NGAKKITITGL-HTGCLQLITKNXK-----IKSAAEIKKGGKIGTVAV 152
179 AGSHNIKLRVWLAAGINPGYVSPQDISQIGADALLSTPPQMPSTLEAGTIFGYCV 238
153 AGSPALFAKVLKSKGLK-----VSDKEG-DVSFVYQSDQLQCVLDKGEVDIAL 202
239 GEPNNQAVFVGIGVPTVTDDELMDOTBEK---VFGVTKQMAEKYPTTYLAATKALIRAA 295
203 GDP-DTEVLKKQYGFKTLANSSTDCKGKMEYCCVAYVSNDIYKHPVAAKTTLAMQKAA 261
296 IWLADNNKKEALIEMLAQOYGVADVEVLAASNG-TEYEKDKRALPDFTFPRHG 354
262 NWV-----QKKEBEVTDIQLNQNVVAGSKGSLNLSNLYTK----- 298
355 ASYPSYSSAV---WYLTQLRRMGMIN 377
299 ---PSYSGAYDSPTVASDLRKIGIILS 322

```
RESULT 13
US-10-282-122A-47676
; Sequence 47676, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Mamado, Carlos
; APPLICANT: Mamado, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreysch, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47676
; LENGTH: 192
; TYPE: PRF
; ORGANISM: Burkholderia cepacia
US-10-282-122A-47676

Query Match      5.7%; Score 140; DB 15; Length 192;
Best Local Similarity 26.8%; Pred. No. 0.00021;
Matches 42; Conservative 32; Mismatches 51; Indels 32; Gaps 4;

QY 42 LKFGKLTDMAPLVAARKGFEEDEGLV-OLEAQAANKVMDRVNGELDGSNMLA 100
DB 37 VAIQYLPITDAAPLVAHNNNGYFASGLTVEQPKLRSWAOLVEALSGQVNVVHLSFM 96
QY 101 PLAAAGVGTAKDIEVPSMGFNGNAITVSNEMWQMKNPILLEGKPVHPITKADVLK 160
DB 97 TLMAR--YGSRAPAKVANNHNVGSLTVAAPD-----DSLRLDGGKTV----- 138
QY 161 VEKYAKGKPFNMAMTFPAGSHNIKRLWYLAAGINP 197
DB 139 -----AIPFWYSIHNVVQDMLRREGGLVP 162

RESULT 14
US-10-156-761-8511
; Sequence 8511, Application US/10156761
; Publication No. US20030119018A1
```

```
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8511
; LENGTH: 883
; TYPE: PRF
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8511

Query Match      4.8%; Score 116.5; DB 14; Length 883;
Best Local Similarity 21.3%; Pred. No. 0.32;
Matches 114; Conservative 62; Mismatches 147; Indels 213; Gaps 35;

QY 6 RSSSKLLITLSASIAWGLTAPDVGA-----VGKLEKEDLKFGPIKLTDMAP-----L 55
DB 15 RSLGRSLALTMTAVAGAGIALSPASADSGPTTADPTFAHGLKGEY--YTQAPGAPDF 72
QY 56 AVAAKGFPEDEGL-FVQLEAQAANKVMDRVNGELDGSNMLAPPLAASVGFGRKADI 114
DB 73 ASIKATGF--DPLDPLDSLSESR-----LSAATGRADVAV 105
QY 115 E-----VP-----FSM-GFNGNAITVSENEI-----W--HQMKPNIPLEGKRVHP 151
DB 106 RWTGEVPEKSGPTTSTIGDNGFRIMVQKLTIDHWUDWDBREQSGEVLDLTAQR-AVD 164
QY 152 IKADYLKPVVEKYKAKGKPFNMAMTFPAGSHNIKRLWYLAAGINGYVSPQ----- 204
DB 165 IKVEYEFHF-----CGSNLHLR-WTPPGSKS---AVPQSAFRLPD 201
QY 205 --DISQIGADAL-----LSVTPEPPQMP-STLEAGTIFGYC 237
DB 202 GPDYNGALSATVLDGGRVRLDPPQALASAPATVLDHLTVTSGKWPVLSVKT----- 254
QY 238 VGEPMNOQAVFGIGVPIITDE-----ELMKD-----TPEKVFQVTKOMARK 279
DB 255 --DPSDARALLVGLGEPVGNKAGTRAKGAVNTYDGKALADTSGTPVAF-----WSSG 307
QY 280 YPN--TYLAVTALIRAAIWLADANNKRNKEAIEMLAOKQYVGADEVILA-ASMGTFEY 336
DB 308 -FNKSTYELRTK-----WADQVGPKN-----ALPEYRPPQLTRKAMQNLNLTWQF 351
QY 337 -----EKODGRALPDFTFFRHGASYP-----SYSAWYLTQLBRKMNINEK 380
DB 352 AGAKAGEEPPRTRTLA-----RIIVPYVESQLSGIEHEDRMWY-----RTFTV----- 398
QY 381 PDNWTYLDTAKV-----YRPDIYLAAXELVAGKAKADFPADT--SIRKPSQ 426
DB 399 PADMKVGHGKRLRLNFGAVDWQSEVTVNGTK--VAEHKGGYRKFSADVTDLAKPGR 452

RESULT 15
US-10-156-761-9969
; Sequence 9969, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
```

APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9969
LENGTH: 353
TYPE: PR1
ORGANISM: Streptomyces avermitilis
US-10-156-761-9969

Query Match 4.6%; Score 111.5; DB 14; Length 353;
Best Local Similarity 18.4%; Pred. No. 0.23;
Matches 71; Conservative 55; Mismatches 161; Indels 99; Gaps 12;

QY 5 IRSSSKLLITLSASLAVWGLT-IAPDVGA-----GKLEKEDLKFGFIKLTDMA 53
DB 1 MRETAATYLAAGLALSSLTACANDASTASSGSGKDGKVKYIMVGGIDKVIYL 60
QY 54 PLAVAAEKGFPEDEGLFVQLEAOANWKVMDRVNNGELDGSHMLAPAPLAASVGF----- 108
DB 61 PAULTQRLGTFDEAGLDVELSEPAQVQATALVSGVQC-----AVGFYDHTL 109
QY 109 -----GTKADIEVPFSGKFNAGNAITVSNELIWHQMKPNIPLGGKPVHPIKADYLKPYVEK 163
DB 110 DLQTKGDVRSVVOFSHA-PGEVEIVSNR-----HADITSP 145
QY 164 YKAEKGFPMNAMPFPAGSHNIKRLRYMLAAGINPGYISPPQDISGQIGADALLSYTPPQ 223
DB 146 KDFKGR-KLGIIGLSSSTDFLTXYLAVKXGVSSEFSP-----VAVGAGPT 190
QY 224 MPSTLEAGTIFGYCVGEPNNOAVFPGIGVPI-----TDEELMKDTPPEKVGVTQKW 276
DB 191 FISALQKGSIDGGMTTDPYANILAKNLGKVLIDMTFRGSKKALGGPYPSSSLYMOTEW 250
QY 277 AEKYPNTYLAVTKALIRAAIWLADANNKRRKALIMLAQK---QYVGADVEVLAASNGT 333
DB 251 VNSHKDTVQKLANAFVYTKLMGT-----HTASQIAEKMPADYSGGNKTLVAGAIKTT 303
QY 334 FEYKXD-----KRALPDFN 348
DB 304 LPMFTDDGVMPKDXPETVERVLKAFN 329

Search completed: May 18, 2005, 17:49:23
Job time: 147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 17:36:01 ; Search time 57 Seconds
(without alignments)
783.238 Million cell updates/sec

Title: US-10-689-200-2

Perfect score: 2443

Sequence: 1 MTTIRSSSKLLTLASL.....AKFAIGLKQQTAVAGKVD 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651	26.6	670	2 S77387	nitrate transport
2	627.5	25.7	470	2 AC3037	hypothetical prote
3	627.5	25.7	471	2 H98248	nitrate-binding pr
4	616.5	25.2	441	2 E95928	probable nitrate t
5	611.5	25.0	431	2 AD3037	hypothetical prote
6	611.5	25.0	431	2 G98248	hypothetical prote
7	610.5	25.0	440	2 AG1882	nitrate transport
8	587	24.0	657	2 A11882	nitrate transport
9	576	23.6	430	2 C95300	Nrta-type periplas
10	572	23.4	443	2 JQ2134	nitrate transport
11	549.5	22.5	430	2 G87324	hypothetical prote
12	545.5	22.3	442	2 S58738	nitrate-binding pr
13	538.5	22.0	667	2 AH2165	bicarbonate transp
14	534.5	21.9	667	2 S75959	nitrate transport
15	525	21.5	446	2 S77389	nitrate transport
16	502.5	20.6	458	2 AF2165	bicarbonate transp
17	494	20.2	402	2 F83422	conserved hypochet
18	476	19.5	659	2 S30893	nrtc protein - Syn
19	473.5	19.4	452	2 S75956	nitrate transport
20	470.5	19.3	437	2 AF2322	nitrate-binding pr
21	444	18.2	450	2 A44751	carotenoid-binding
22	398.5	16.3	332	2 S52248	nitrate and nitrit
23	390.5	16.0	426	2 AF2663	hypothetical prote
24	390.5	16.0	426	2 E97445	nitrate transport
25	355	14.5	625	2 S58789	ATP binding protei
26	349	10.2	235	2 B55581	naef protein - Kle
27	240	9.8	339	2 AB3354	hypothetical prote
28	125	5.1	337	2 B75370	probable ABC trans
29	124.5	5.1	333	2 AB3080	ABC transporter, s

30	124.5	5.1	333	2 G98206	hypothetical prote
31	116.5	4.8	468	2 AB3359	hypothetical prote
32	114	4.7	325	2 H75414	probable ABC trans
33	114	4.7	336	2 AB3801	ABC transporter (s
34	111	4.5	310	2 AC3238	hypothetical prote
35	109.5	4.5	300	2 H69260	thiamin biosynthes
36	106.5	4.4	675	2 D48492	kpac protein - Bsc
37	106	4.3	483	2 G86811	sugar ABC transport
38	106	4.3	1273	2 T34558	hypothetical prote
39	105.5	4.3	339	2 E95962	probable taurine u
40	105.5	4.3	566	2 A40589	cellulase (EC 3.2.
41	104.5	4.3	1135	2 T14659	DNA polymerase III
42	104.5	4.3	1174	2 T15021	probable DNA polym
43	103	4.2	897	2 AC2001	hypothetical prote
44	102	4.2	301	2 B95947	phosphate uptake A
45	102	4.2	1291	2 I40631	non- proteolytic bo

ALIGNMENTS

RESULT 1

S77387 nitrate transport protein C-2 - Synechocystis sp. (strain PCC 6803)

N/Alternate names: protein s11452

C/Species: Synechocystis sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004

C/Accession: S77387

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S77387

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-670 <KAN>

A/Cross-references: UNIPROT:P73450; EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BAA1749(C

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Genetics:

A/Gene: nrtc-2

C/Superfamily: ATP-binding cassette homology

C/Keywords: ATP; nucleotide binding; P-loop

P/25-215/Domain: ATP-binding cassette homology <ABC>

F/42-49/Region: nucleotide-binding motif A (P-loop)

Query Match 26.6%; Score 651; DB 2; Length 670;

Best Local Similarity 36.3%; Pred. No. 9.7e-40;

Matches 158; Conservative 68; Mismatches 153; Indels 56; Gaps 13;

QY	28	ADPVGAVGK-----LEKEDLKGFILKLTMAFLVAAGKFPEDDEL-FVQLEAQA 77
DB	260	AKKVGAVSOFPAEMGANGLEKINLIDGFIPLTDCAPLVVAKXGFPQKLEOVNLVKEP 319
QY	78	NMKVMDRVYNGELDGSNMLAPAPLAASVFGTKADIEV--PSMGNGNAITVSNIEIWH 135
DB	320	SWQAIADGIRERRLDGQWAGMPALITLGMGSKTPIPMVTANVSRNGAIIITSKKFAE 379
QY	136	QMKNPILBEGKGVPHIKADYLLKPVVEKYKAEKPPRMAMTTPAGSHNIRLRYWLAAGI 195
DB	380	-----AGVKTLDELRL-----LKLAEITPPQVSTLGVHVASMQNLLLRWLASGSI 424
QY	196	NPGYISPPQDISQIQADALLSTPPQNPSTLEAGTIFGYCYGEPNQQAIVKGIQVIV 255
DB	425	D-----PDDQIN-----LKRLLPPOWWSNLEAGNIDGFCGEBPWSYAVAKQNGYVI 471
QY	256	ITDEELMKDTPPEVPGVTKQMAEKYENPTYLAATKALIRAIIMDADNNKRAKAIEMLAQ 315
DB	472	ATDLIDIMNGHPEVYLVGRREWVVKIPATHLALVKALLBACEY--CDRRRKHQIILYIAL 529
QY	316	KQYVGADVEVLAASNGTPEYERD-DKRALPDENTFRRGASYSYSSAVVYLTOLRRWG 374

A:Gene: nrCA, SMD21114
A:Genome: plasmid
S:Galibert, F.; Ffman, T.M.; Long, S.R.; Pubher, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chah, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenhof, M.; Votholler, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: nrCA, SMD21114
A:Genome: plasmid

Query Match	25.2%;	Score	616.5;	DB 2;	Length	441;			
Best Local Similarity	33.8%;	Pred. No.	1.8e-37;						
Matches	153;	Conservative	74;	Mismatches	182;	Indels	43;	Gaps	12

QY		19	SLAWGTLIADVAVGLLEKEDLKFGIKTIDNAPLATAAEKEGFPEDEGFLVTOEAQN	78
Db		19	NLSAGGALLPAP--GSIRHESGRIVRAGTIPLVDAAVLIAAEPFAQREGITTELIVDVS	76
QY		79	KWVMVDRVNGELDGSHMLADAPLAASVGFTGTADI EV-PSMGFNGNAITSVEINHWOM	137
Db		77	WANVRDLAFRFQFVAHMLSMPVAAMLGLSINSPTIAPSPISRGGNAILTILSTRLYGLM	136
QY		138	KPNIPLEGSKRVNHIIKADYLKPVVEKYKAEBK-PENNMATPPASHNTKLPMTYLAAGIN	196
Db		137	QQAGCLGGGEBAAL-KAKALAAVIRISAGAACKPPLTGVTYPSSHNEFFMYTLAAGID	195
QY		197	PGYYSPEODISSGIADALLSVTPPPQMSTLEAGTIFGVCGEPMNQOAFKIGVPVI	256
Db		196	P-----DRIYKL VVVP PPMTSDBLLAAGALDGF CVGA PMNVVASREGVRI VA	242
QY		257	TDEELMKOTPEKVHVGTKQMAEKY PNTYLAATKALIRAIYLDADNNKNRKREALIMLAOK	316
Db		243	TKODIWPSSAPEKTVIGMRPEWAENARDIVSRVLVLDRAARV--SDBPANRGRLAEVIAEE	300
QY		317	QYVADVPEVLAAASNUGTFEYEKD-DKRALPDNFNFRHGASGPSSAYVWLJTLBRGM	375
Db		301	RHAVAAPPDIIRVULAGEPAIDPEGNRKRVENEYLVFHAGFAPVPRPSQALMTYISQWRKQ	360
QY		376	INEFKPPDNWYLDTAKNVYRPDIYLA AEKELVAEBKAKADEPADTISI-----KPSONFI	430
Db		361	TSLSRRQ---RVDAAISAYRPDL YREAL-----GK---DALPADADAGLEGATGDGRFM	408
QY		431	DKVFPDANKRPNDYLAKRALGLKGQTYAGKCV	462
Db		409	DGHVFDPGRLEDYIKSF-----GAPS VNSTRV	435

RESULT 5
AD3037
hypochemical protein nrtc {imported} - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AD3037
R:Wood, D.W.; Seibthal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCiell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
eter, E. W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD3037
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431 <KUR>
A:Cross-references: UNIPROT:O8U924; GB:AB006689; PIDN:AAL44714.1; PID:g17742345; GSPDB:C
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: nrtc
A:Map position: linear chromosome

Query Match	25.0%;	Score 611.5;	DB 2;	Length 431;
Best Local Similarity	35.5%;	Pred. No. 4.1e-37;		
Matches 151;	Conservative 69;	Mismatches 166;	Indels 39;	Gaps 12;

```

QY      34 VGEKKEBOLKRFQFILTDTMAPLAVAAEKKFFBEDEGFEVQLEAOANWYKMDRVVNGEJLDG  93
Db      24 VGSBRQKILTRACFILTVDASVLIANAABFADREGILDLVKVSWANVDRDLAERODI  83
QY      94 SHMLAPABLAASVGFQYADIEV-PEMSGNGNAITVSNIEIHWQKNIPL-EGGKVPYH  151
Db      84 AHMLSPMEVASWLGSGNPSPTITPFSILRGGNAITLSTRLPARKALTGISETAGALEN  143
QY      152 IKADYLKRVVEYKKAEGK-PEFMAMTFRAGSHNIKURVYLAAGINGYSSPPDISQI  210
Db      144 ARA--LKLVLDMRARGEARPTLGMTPPSSSHYERRYWLAAGIHDPH-----  190
QY      211 GADALLSTPPQOMSTLEAGTIFGYCVGEPMNQAVFKIGIVITDEELMKDTPKEVF  270
Db      191 --DYKLVVPPPMTSDALAAGIDGFCVGAHPNIIYVAABRGVGIYAAKODLMPAPERYI  248
QY      271 GVTQMAEKYENTYLAVTKALIRAAIWD-ADNNKRKEALIEMLAQOQYVGADEVILAAS  329
Db      249 GMRPEMAEQOETVGRLLTALDAASWCDLADHNDALSGA--LADPRYICAPOSITIRV  305
QY      330 MNGTFPEYE-KODKRALPDENTFFRHGASYPYSASAWYLTQLARKGKINEFKPDNWIYLT  388
Db      306 LAGEFSIDSGRRRIYEKFTYFHGDHANYPROSQSLMYSQMIRWG--OAEISETGVNA  362
QY      389 AKNVYRPDIYLAAYELVABGAKAADFPADPSIK-----PSQNFIDIKVFPDANKPNYD,  444
Db      363 ALSAVRPDIYRAA---LDGKA-----PGDADIRLEGDEGRFVUDGPFVDPADDIAGYV  413

```

OY 445 AKFAI 449
 ||:
 Db 414 NSFAV 418

 RESULT 6
 G988248
 hypothetical protein AGR_L1881GL [imported] - Agrobacterium tumefaciens (strain C58, Cerevisiae)
 C/Species: Agrobacterium tumefaciens
 C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C/Accession: G988248
 R/Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2326, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 ;Reference number: A97359; WUID:21608551; PMID:11743194

[illegible]

```

Qy 211 GADALLSVPPQMPSTLEAGTIGYCVGEPNQAQVAFKGIQVPIITDEBELMKDTPKVF 270
Db 191 --DVKLVPPEPTSDALAGIDGFCVGPVNIIVAAERVGRIIVAKQDLWMSAPKVI 248
Qy 271 GVTQNAEKYPTNYLAVTKALIRAIWLD-ADNNKKNKEAIEMLAQKQYGVADVEVLAAS 329
Db 249 GMRPENAESQGETVGRLLTLMDAAASWCDLADNHDLSGA---LADPRYICAPQSIIRRV 305
Qy 330 MNGTFEYE-KODYRALPDENYTPFRHGASYSYSAVWYLTQLRMGKINEFKPDNYLDT 388
Db 306 LADEFIDISQGNRRVIEKYFTFGDHANYPROSQSLIYQIMRWG---QAEISEGVNA 362
Qy 389 AKVVPDYIYLAAKELVAGSKAKEDFPADTSIK---PSQNFIDKVPEDANKENDYL 444
Db 363 ALSAIPDIYRAA---LGDGKA---PEDADIRLEGDEGRFVDGFEVDPAIAGYV 413
Qy 445 AKFAI 449
Db 414 NSPAV 418

RESULT 7
AGI882
nitrate transport nitrate-binding protein nrtA [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AGI882
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807, M01D:21595285, PMID:11759840
A/Accession: AGI882
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-440 <KUR>
A/Cross-references: UNIPROT:Q44292; GB:BA000019; P1DN:BA072566.1; P1D:G17129954; GSPDB:G
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: nrtA

Query Match 25.0%; Score 610.5; DB 2; Length 440;
Best Local Similarity 33.4%; Pred. No. 5e-37;
Matches 158; Conservative 71; Mismatches 173; Indels 71; Gaps 17;

Qy 8 SSKKLLLT-----LSASIAVWGLTI-----APDVGAVG--KLEKEDLKF 44
Db 5 SRKKLFTTGAAMAAASILVHGCTSNQSOSATGEOAPSAAPAAVNAAPKVEITKALD 64
Qy 45 GRIKLTMAPLVAAEKGFPEDEGLF-VOLEAQANKVVMDRV---VNGELDGSHTLAP 99
Db 65 GEIPLTDAPLIIAEKGFPAKGMTDIEVIKQKSPVTRDNLKIGSSGGIDGAIHILSP 124
Qy 100 APLAASVSGGTADIEVPSMGNGNAITVSNIEHOMKPNIPLEGKRVHPKADYLRP 159
Db 125 MEYLMNT--NDKVPMTIARLNTNGOASIVAEKF--KELNVLE-----SKSLKD 170
Qy 160 VVEKYKAEKGFPMNMTFPAGSHNIIKLRYWLAAGINPGYSPPOISQIGADALLSVT 219
Db 171 AAIKAKADKKALKMGITFGTHDLMRKWTYLAAGINP-----DQDVLEAV 217
Qy 220 PPPQMPSTLEAGTIGYCVGEPNQAQVAFKGIQVPIITDEBELMKDTPKVFQYTKQMAEK 279
Db 218 PPPQMAANKVNTVVGFCVGEPEWNAQLVNOKIGYSALVYGEIWKDHPKAFSMDQWIBQ 277
Qy 280 VNTYTLAVTKALIRAIWLDADNNKKNKEAIEMLAQKQYGVADVEVLAASMNGTFEYED 339
Db 278 NNNAQAAILMALLEAQOM--CQAEKKEKMKCISDRKYFNVAADIIIRAGKNDY--G 333
Qy 340 DGRALPDF---NTFFHGASYSYSAVWYLTQLRMGKINEFKPDNYLDTAKVYRPD 396

```

```

Db 334 DGRKQNFARHMKFADNMSYPYKSHDIWFLTEDIMWGLPK---DTKQDIYNQVNKED 390
Qy 397 IYLAAKELVAGSKAKAEDFPADTSIKPSQNFIDKVPEDANKENDYLAKFAI 449
Db 391 LMKKAAR---AIGVADAE-IPASS--RGVETFF-DGVKEDPEKPEEYLSIKI 437

RESULT 8
AGI882
nitrate transport ATP-binding protein nrtC [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AGI882
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807, M01D:21595285, PMID:11759840
A/Accession: AGI882
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-657 <KUR>
A/Cross-references: UNIPROT:Q9Y276; GB:BA000019; P1DN:BA072568.1; P1D:G17129956; GSPDB:G
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: nrtC
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 24.0%; Score 587; DB 2; Length 657;
Best Local Similarity 35.8%; Pred. No. 4.6e-35;
Matches 150; Conservative 74; Mismatches 139; Indels 56; Gaps 17;

Qy 39 KEDLKGFIKLTDMAPLVAAEKGFPEDEGL-FVOLEAQANKVVMDRVNGELDGSHTL 97
Db 276 KAVIEIGFPLTDSAPLIVAEKGFPAKYGILNVILNANMQALATGVYTGILDAAGV 335
Qy 98 AAPPLAASVSGGTADIEV--PFSMGFNCAITVSNIEHOMKPNIPLEGKRVHPKAD 155
Db 336 AGMPPLTITLGAOSQPTPTYNALNLSRNNAITFSRILYNOGVRSL-----AD 383
Qy 156 YLKPVEKYKAEKGFPMNMTFPAGSHNIIKLRYWLAAGINPGYSPPOISQIGADAL 215
Db 384 -LKAVID--SSPDQILITLVSHASKNQILRYWLAAGID-----PDQVDS----- 427
Qy 216 LSVTPPQMPSTLEAGTIGYCVGEPNQAQVAFKGIQVPIITDEBELMKDTPKVFQYTKQ 275
Db 428 LVYIPTQVWSQLKAGNIDGYCAGPEWNYQAVHDDLGFVAATALIEIWSGPKKVLGVRED 487
Qy 276 MAEKIPNTYLAATKALIRAIWLDADNNKKNKEAIEMLAQKQYGVADVEVLAASMNGTF 334
Db 488 MAQKPEYILNIVKALIEACKY--CDLJNRREIIEILCRPEY---LDVNPAYVRSFPI 541
Qy 335 -EYBKDD---KRALDPNTEFFHGASYSYSAVWYLTQLRMGKINEFKPDNYLDTAK 390
Db 542 DRYDGDGTTPQOLATRYNOFYLNKTYNPIRITLMTWMAWG-LTFP--PKWV--VEITR 598
Qy 391 NVYRPDIYLAAKELVAGSKAKAEDFPADTSIKPSQNFIDKVPEDANKENDYLAKFAI 449
Db 599 RVCRTDIFFAARDL-----GLLDIGDDPI-----HLFDGTL-FNPSPIYILASLEI 646

RESULT 9
G95300
NrtA-type periplasmic nitrate transport binding protein, probable Sma0585 [imported] - Sj
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: G95300
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowne
J.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A/Reference number: A95262; M01D:21396509; PMID:11481432

```


Best Local Similarity 32.7%; Pred. No. 1,4e-32;
Matches 148; Conservative 68; Mismatches 179; Indels 57; Gaps 16;

QY 6 RSSSKLLTLTSLAIVM-GLTTPDVGAAGLEKEDLKFGFIKLTDMAPLAAAEKGF 64
DB 16 RSLMGAATLAAAKAFPGAHAAAGP-----EVAKARLGFIATLDSPLIATKERGLF 71
QY 65 EDEGL-FVQLEAQNKKVMDRVY-----NGEIDGSHMLAPAPLAASVGF---GTAKDIEV 116
DB 72 AKKGMDIEVYKQASAAATRDNLVLAERGDGAHILTPMPYLMTGTITGAPTPMYI 131
QY 117 PPSMGNGAIVTSNEIWHQMKENIPLEGKPVHPKADYLKPEVKEKYAEKPFMMAT 176
DB 132 LARLNTGGISIGNDL-KSVKGLNSAGA-----KAKEL-----QMKAGIAIVANT 179
QY 177 FRAGSNIKLRWYLAAGINPGYSPPODISQIGADALLSTPPQMPSTLEAGTIFGY 236
DB 180 FRGTGTHLIRWYLAAGINP-----DADVSTIPIPPQMANMKAAGTQDAF 226
QY 237 CGGEPNQAQAVFPGIGVPIITDEELMKDPEKVGVTQMAEKYPTMYLAVTKALIRAAI 296
DB 227 CGGEPNQAQAVFPGIGVPIITDEELMKDPEKVGVTQMAEKYPTMYLAVTKALIRAAI 296
QY 297 WLDADNNKRAEIAEMIAQKQYVADVEVLAASMGTFEYKDKRAL---PDPTFFPH 353
DB 287 W--CDKAAVLPGQCSIVSGKQYVNVPMGDLPRLOGTVDY--GDGRTLNKSPHRMKTAD 342
QY 354 GASYPSSAAYVYLTQLRWGMINEFKPDNWLDTAKNVPDIYLAAKEVAESKAKA 413
DB 343 NASFPKSHDLWFLTEDIRMGVLPQKTNKALVD--KNVRSDIMRAAKSIGSG---- 395
QY 414 EDFPADTSIKPSQNFIDKVPDPANKPNXYLA 445
DB 396 ---PAGDS-RGVERFPDGKV-FDPANPAGYLA 422

RESULT 12
558738
nitrate-binding protein nrtA precursor, periplasmic [similarity] - Phormidium lamosum
C/Species: Phormidium lamosum
C/Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C/Accession: S58738; S56641; S62124
R/Merchan, F.; Friebo, R.; Kindle, K.L.; Liama, M.J.; Serra, J.L.; Fernandez, E.
Plant Mol. Biol. 27, 1037-1042, 1995
A/Title: Isolation, sequence and expression in Escherichia coli of the nitrite reductase
A/Reference number: S56640; MUID:95284340; PMID:7766873
A/Accession: S56641
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-67, 'A', '69-442 <MEW>
A/Cross-references: EMBL:Z19598
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C/Genetics:
A/Gene: nrtA
C/Keywords: nitrate transport; periplasmic space

Query Match 22.3%; Score 545.5; DB 2; Length 442;
Best Local Similarity 33.3%; Pred. No. 2,9e-32;
Matches 145; Conservative 67; Mismatches 175; Indels 49; Gaps 14;

QY 27 IAP-DVGAVGKLEKEDLKFGFIKLTDMAPLAAAEKGFEDGL-FVQLEAQNKKVMD 84
DB 44 IAPAVAAADTPEVTTAKLGFICPDGLALIIAKEKGLPAKYGMPDVEVLKQASAAVTRD 103

QY 85 RVV-----NGEIDGSHMLAPAPLAASVFGTKADIEVPS-----MGFNGAIVTSNEIWHQ 136
DB 104 NIELSGGGGIDGAHILTPMPYLSAGTITKQKQKVPKNIARLANTNGGICLANTY--- 160
QY 137 MKRNIPLLEGKAPVHPKADYLKPEVKEKYAEKPFMMATFPAGSHNIKLRWYLAAGIN 196
DB 161 ----KDLKIGTSSPLKEAFKA-----KAEKEIQAATVPFGTHTDLMWRWLSAGID 211
QY 197 PGYSPPODISQIGADALLSTPPQMPSTLEAGTIFGYCGGEPNQAQAVFPGIGVPI 256
DB 212 -----PDKDIS-----TIVPPQMANIKVNMMEFECGEPWPAQTQVNGGLGYNAM 258
QY 257 TDEELMKDPEKVGVTQMAEKYPTMYLAVTKALIRAAIWADANNKRAEIAEMIAQ 316
DB 259 TTGELMKDPEKAFARADWQHPKAAVALMAVQEAQIW--CDPAPKEKQCVISGR 316
QY 317 QYVADVEVLAASMGTFEYKDKRALPDPTFFHGASYSYSASVAVYLTQLRWGMI 376
DB 317 EMFKVVEDILERSKGNFLGVRQLENSPLMKFMNDASYPFKSHDLWFLTEDIRMGYL 376
QY 377 NEFKPDNWLDTAKNVPDIYLAAKEVAESKAKAEDFPADTSIKPSQNFIDKVPDI 436
DB 377 PADDTKALVDA---VNREDLWREAKAIGQEA---IPASTS-RGVETFEF-DGVKPD 426
QY 437 ANKPDNYLAKFALGLK 452
DB 427 PENPSAYLS--ALKIK 440

RESULT 13
AH2165
bicarbonate transport ATP-binding protein cmcC [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AH2165
R/Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, S.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AH2165
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-667 <KUR>
A/Cross-references: UNIPROT:O8YT47; GB:BA000019; PIDN:BA074578.1; PID:G17131973; GSPDB:G17131973
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: cmcC
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 22.0%; Score 538.5; DB 2; Length 667;
Best Local Similarity 33.3%; Pred. No. 1,7e-31;
Matches 140; Conservative 70; Mismatches 158; Indels 53; Gaps 14;

QY 30 DVGAVGKLEKEDLKFGFIKLTDMAPLAAAEKGFEDGL-FVQLEAQNKKVMDRVYN 88
DB 270 DVARRG-LEKWLLEGLFLPTACAPLAAAEKGFETKGLDENVLVRETSMPGIVDGKMG 328
QY 89 GELDSHMLAPAPLAASVFGTKADIEV--PPSMGNGAIVTSNEIWHQMKENIPLEGK 146
DB 328 GTYDAQNPSSGMPMLTTLGHDNQLPVYVLTALTRNGAIIATLAKRPD-----EGV 380
QY 147 KPHADIKADYLKPEVKEKYAEKPFMMATFPAGSHNIKLRWYLAAGINPGYSPPODI 206
DB 381 RLDSLPK-NYL-----LRTDQRIHMGVYHPSAHNILLRKLWLAAGIDP----- 424
QY 207 SCQIGADALLSTPPQMPSTLEAGTIFGYCGGEPNQAQAVFPGIGVPIITDEELMKDTP 266
DB 425 ----DLDVDMRTIPPAQWADLQNSIDGYCGGEPNRAAIVENGFITATDLEWVLGHP 480
QY 267 EKVPQVTKQMEKYNTYLAVTKALIRAAIWADANNKRAEIAEMIAQKQYVADVEVL 326

Db 481 GKVLGVRBMAERYEPTHTIALTKALLBACY--CSPRENVEVRRIVAGDVIYSTDLDYI 538

Qy 327 AASMNGTFEYKDKRALPDF--NTEFRHGA-SYPSYSAVWYLTOLRWMGINEFKDN 383

Db 539 QLEDPNSLVICDDI--HPLADYAHGHFFAESAINRPTBQIWMISQIARNG--DTFPPRN 594

Qy 384 WYLDIAKANNVYRPDIYLAALKEVLAEGRKAKAEDFPADTSIKPSONFPIDKYPPDANKNDY 443

Db 595 W-VEEVERVCVRVVFSTAAREL-----GLDISYTRQPIELFDGTFPNADDPPIAY 642

Qy 444 L 444

Db 643 L 643

RESULT 14

S75959

nitrate transport protein C-1 - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: protein slr10043

C:Species: *Synechocystis* sp.

A:Variatey: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004

C:Accession: S75959

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

S:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75959

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-667 <KAN>

A:Cross-references: UNIPROT:O55462; EMBL:D64006; GB:AB001339; NID:91001291; PIDN:BAAL080

C:Genetics:

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 21.9%; Score 534.5; DB 2; Length 667;

Best Local Similarity 31.7%; Pred. No. 3.3e-31;

Matches 146; Conservative 73; Mismatches 154; Indels 87; Gaps 16;

Qy 10 KULLTLTSLAVWGLTTPDVGAVGLKEDLKFGFIKLTMDAPLAVAEKGFEDGL 69

Db 261 KGLRAKKTALIAHGH-----LKKVNLDELGYPLVACAPLVVAKGFFAKHGI 308

Qy 70 -FVLEAQAAMKVVMDRVVNGELDGSMLAPPLAASVGFSTKADIEV--PFGMGFNQA 126

Db 309 DEVSLVRETSKRGIVDGIAGYIDGAQMPAGMFTWLTAGYRQSIIPVVSALTMTKNA 368

Qy 127 ITVSNIEIWMQKNPIPLBEGKPVHPRIKADYLKVEVKYAEGRPFMMAMTFPAGSHNIX 186

Db 369 ITLSKLYDQ-----GIYTAE-DPRQLL---ASDGRHTLGMVHPMSMNL 413

Qy 187 RTWLAAGGINPGYSPDISQIGADALLSTVPPQMFSTLEAGTIFGVCVEPNNQA 246

Db 414 RTWLAHNNP-----DRDVLKTIIPQAWADLAKGITDGCVSSEPNLRA 460

Qy 247 VKGIGVPIVTDELMKDTPEKVGVTAKMAEKYPTVYLAVTALRAAIWLDADNNKOR 306

Db 461 SNEGAGFSLATDLEIQQNHPGKVLGVRBDYALHPHVAHYALAEACAYC-ADPN-HE 518

Qy 307 KEAIEMLAQOYVADVEVL-----AASMNGTFEYKDKRALPDFTFRHGASYP 358

Db 519 MEIRBELLATROYLSTINDYIHLGDPRGRCRLGNPEY-----SHLFFGQGFNP 569

Qy 359 SYSSAVWYLTOLRWMGINEFKDNWYLTDAKANNVYRPDIYLAALKEV-----V 406

Db 570 SRTHELMNMTQAWMGDIP--PFRNW-VEILERVCGVGFSTAARELGYDVNYORQPIAL 626

Qy 407 AEGKAKAEDFPADTSIK-----PSQNFIDKVPFANKP 440

Db 627 FDGKV-----FNADDPYALINQVTIHRNFTIAEVLNPTP 662

RESULT 15

S77389

nitrate transport protein A-2 - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: nitrate transport 45K protein; protein slr1450

C:Species: *Synechocystis* sp.

A:Variatey: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S77389

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

S:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S77389

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-446 <KAN>

A:Cross-references: UNIPROT:P73452; EMBL:D90906; GB:AB001339; NID:91652492; PIDN:BAAL7492

C:Genetics:

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

A:Gene: nrcA-2

Query Match 21.5%; Score 525; DB 2; Length 446;

Best Local Similarity 28.5%; Pred. No. 9.3e-31;

Matches 139; Conservative 78; Mismatches 186; Indels 84; Gaps 15;

Qy 1 MKTIISSSKKULLTSLASIAVWGL-----TIAPVGAVGKL 37

Db 1 MSNFSRSTKRKEFTYGAALAGGVVLHGTSPTTSTGTGSSSTDQALSPVBEENAP 59

Qy 38 EKEDLKFGFIKLTMDAPLAVAEKGFEDGL-FVLEAQAAMKVVMDRVV-----NGELD 92

Db 60 EYVTLALGFIALTDAPLIIAEKGFYAKGMVDVLAQASGTTIRDNVLGASAGSID 119

Qy 93 GSHMLAPPLAASVGF--GTKADIEVPPSMGFNGNAITVSNIEIWMQKNPIPLBEGKPV 149

Db 120 GAHILTPMYLYLTGMVTGDKPTPYILARINVGQIQLGNNY-----KDLKVGTD 172

Qy 150 HPYKADYLKVEVKYAEGRPFMMAMTFPAGSHNIXRTWLAAGGINPGYSPDISQ 209

Db 173 APLKEAFATVDPK-----VAMTFPGGTHDMIRYWLAAAGMEP----- 211

Qy 210 IGADALLSTVPPQMFSTLEAGTIFGVCVEPNNQAVFGIGVPIVTDELMKDTPEKV 269

Db 212 -GNDPSTIVPQAWANVKNAMBSFCGEPPPLDTVAGVGYALTTGQMKDHEKA 270

Qy 270 FGVTKQMAEKYPTVYLAVTALRAAIWLDADNNKKEAIEMLAQOYVADVEVLAAS 329

Db 271 FGRADWVQNPRAKALMAVMAEQM--CDQAKKEKMCQILSKREMPKVPFEDIIDR 328

Qy 330 MNGTFEYK-----DKRALPDFNTFRRGASPSYSSAVWYLTOLRWMGINEFKDN 364

Db 329 SKGIYNGGQETFEDEIM--QKYWVNASAPYSHQWFLTENIRWGYLAASDTTKA 385

Qy 385 YLDIAKANNVYRPDIYLAALKEVLAEGRKAKAEDFPADTSIKPSONF--FIDKVPDANKPN 442

Db 386 IYD--KVRREDIMRELAQAL-----EVPADQIPS-----SPSRKRIETFFDQITDPPENPQA 434

Qy 443 YLAKFAI 449

Db 435 YLDSLKI 441

Search completed: May 18, 2005, 17:43:45

Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 17:36:47 ; Search time 179 Seconds
(without alignments)
1327.401 Million cell updates/sec

Title: US-10-689-200-2

Perfect score: 2443
Sequence: 1 MKTIRSSSKLLTLISASL.....AKFALGKQGVAGKVD 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1597	65.4	462	2	Q8KY55 azospirillum
2	1586	64.9	475	2	Q7MDT5 Q7mdt5 vibrio vuln
3	1572	64.3	487	2	Q8D6X8 Q8d6x8 vibrio para
4	1564	64.0	454	2	Q87HA7 Q87ha7 vibrio para
5	1257.5	51.5	553	2	Q7UYV1 Q7uyv1 rhodospirillum
6	651	26.6	670	1	NRTC_SYNY3
7	627.5	25.7	430	2	Q8U925 Q8u925 agrobacteri
8	616.5	25.2	471	2	Q7CTL1 Q7ctl1 agrobacteri
9	616.5	25.2	441	2	Q92VK8 Q92vk8 rhizobium m
10	611.5	25.0	431	2	Q8U924 Q8u924 agrobacteri
11	610.5	25.0	440	1	NRTA_ANASP
12	610	25.0	429	2	Q7NEC4 Q7nec4 anabaena sp
13	594	24.3	427	2	Q7NKB1 Q7nkb1 anabaena sp
14	587	24.0	457	2	Q8Y2F6 Q8y2f6 anabaena sp
15	585.5	24.0	420	2	Q8Y2F5 Q8y2f5 anabaena sp
16	582.5	23.8	439	2	Q898C9 Q898c9 bradyrhizob
17	578.5	23.7	412	2	Q98H14 Q98h14 rhizobium l
18	578	23.6	420	2	Q6D2V0 Q6d2v0 sinorhizobium
19	577	23.6	439	2	Q8D7J8 Q8d7j8 sinorhizobium
20	576	23.6	430	2	Q922Z1 Q922z1 rhizobium m
21	575.5	23.6	418	2	Q48466 Q48466 klebsiella
22	572	23.4	443	1	NRTA_SYNP7
23	567	23.2	454	2	Q8D7F6 Q8d7f6 synecococc
24	566	23.2	425	2	Q98H12 Q98h12 rhizobium l
25	559.5	22.9	388	2	Q89LH2 Q89lh2 bradyrhizob
26	549.5	22.5	430	2	Q9AA18 Q9aa18 caulobacter
27	548.5	22.5	448	2	Q7N1U6 Q7n1u6 gloeobacter
28	548.5	22.5	458	2	Q8DH7F Q8dh7f synecococc
29	548	22.4	428	2	Q6WRT2 Q6wrt2 rhodobacter
30	545.5	22.3	442	2	Q51880 Q51880 photobacter
31	542.5	22.2	666	2	Q7N1U3 Q7n1u3 gloeobacter

32	540	22.1	655	2	Q7NKA9 Q7nka9 gloeobacter
33	538.5	22.0	667	2	Q8Y147 Q8y147 anabaena sp
34	536.5	22.0	663	2	Q55107 Q55107 synecococc
35	534.5	21.9	667	2	Q55462 Q55462 synecocyst
36	531	21.7	404	2	Q88141 Q88141 pseudomonas
37	525.5	21.5	462	2	Q6N722 Q6n722 rhodopseudo
38	525	21.5	403	2	Q883P3 Q883p3 pseudomonas
39	525	21.5	446	1	NRTA_SYNY3
40	521.5	21.3	533	2	Q6W1F8 Q6w1f8 synecococc
41	510.5	20.9	459	2	Q891A5 Q891a5 bradyrhizob
42	509.5	20.9	559	2	Q7U3E5 Q7u3e5 synecococc
43	506	20.7	386	2	Q6N3I6 Q6n3i6 rhodopseudo
44	502.5	20.6	458	2	Q8Y149 Q8y149 anabaena sp
45	494	20.2	402	2	Q912V6 Q912v6 pseudomonas

ALIGNMENTS

RESULT 1	ID	Q8KY55	PRELIMINARY;	PRT;	462 AA.
AC	Q8KY55;				
DT	01-OCT-2002 (TREMBlrel. 22, Created)				
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)				
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)				
DE	Nitrate transporter substrate-binding protein.				
GN	Name=nsaf;				
OS	Azospirillum brasilense.				
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;				
OC	Rhodospirillaceae; Azospirillum.				
OX	NCBI_TaxID=192;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=59245;				
RA	Screenoud O., Placek D., Verreth C., Vanderleyden J.;				
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF283498; AAM73543.1; -				
DR	InterPro; IPR002110; ANK.				
DR	PRINTS; PR01415; ANKYRIN.				
SQ	SEQUENCE 462 AA; 50270 MW; 25C86DP070421EBC CRC64;				
Query Match	65.4%; Score 1597; DB 2; Length 462;				
Best Local Similarity	65.6%; Pred. No. 6.3e-112;				
Matches 301; Conservative 60; Mismatches 92; Indels 6; Gaps 4;					
QY	SSSKLLTLISASLAWGLTAPDVAGVGLEREDLKFGITKLTMDAPLAVAEKGFED 66				
DB	6 ASPTALISAAATLA--LMLGSAQAAPLVEKDQKLGIFKLTMDAPLAIAEKGFED 62				
QY	EGLFVQLEAQAAMKVMDRVVNGELDGSMLAPAPLAASVGFQTKADIEVPSMGFGNA 126				
DB	63 EGLSVTLLEPQANKKVLIDRYISGELDGAHMLAQPIGATIGFTQANVVTAFSMDLNGK 122				
QY	127 ITVSNIEIWMQKKNIPPL-EGKPFVHPKADYLPVVEKYKAEKPPNMAWTFPAGSHIK 185				
DB	123 ITLSNVEWMKKNLPKPGDKPIHPIKADALPVAIAQRAEKKPFTMGVFPVSTHNYE 182				
QY	186 LRYWLAAGGIPNGYSPPODISQIAGDALLSTPPPPMPSTLEAGTIFCYCGEPNQQ 245				
DB	183 LRYWLAAGGIPNGYSA-PDVDSQIQADALLSTPPPPMPATEAGTIFCYSGEPNQQ 241				
QY	246 AVFGKIGVPIITDEELMKDTPKVPFVNTQMAKYENNTYLAVKALIRAIWLDADNNK 305				
DB	242 AVFGKIGVPIITDEELMKDTPKVPFVNTQMAKYENNTYLAVKALIRAIWLDADNNK 301				
QY	306 RKEAIEMLAQKQYVADVETLAASNNGTPEYKDKRALPDENTFPHGASYPSSAAW 365				
DB	302 RAAVAVILAKSEVYVADAVIANSMGTFEYKGRADVDFVFRVYNAVTFYSDAVW 361				
QY	366 YLTQLRKGKINLFFKPDNWTLDPAKVYRPDIYLAAKSLVAGKKAEDFP-ADTSIYP 424				
DB	362 YLTQMRWQOIAEAKPDADYDEYARVYRREIYLKARLLVEGKAKEADFPWTSQGYX 421				

QY 425 SONFIDKVPFDANKPNDYAKAIGIKKQTVAGKGV 463
 DB 422 LDNGFIDGAYDGRKRENYLTKLPIGLKGQAVQGGQLV 460

RESULT 2

Q7MDT5 PRELIMINARY; PRT; 475 AA.
 AC Q7MDT5;
 DT 01-MAR-2004 (Tremblrel. 26, Created)
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
 DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Nitrate transporter system, periplasmic component.
 GN OrderedLocusNames=VVA0951;
 OS Vibrio vulnificus (strain YJ016).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxId=196600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14656965; DOI=10.1101/gr.1295503;
 RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
 RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
 RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
 RT "Comparative genome analysis of Vibrio vulnificus, a marine
 pathogen.";
 RL Genome Res. 13:2577-2587(2003).
 DR EMBL; AF005348; BAC96977.1; -.
 KW Complete proteome.
 SQ SEQUENCE 475 AA; 53801 MW; B616BD1915DC3C67 CRC64;

Query Match 64.3%; Score 1586; DB 2; Length 475;
 Best Local Similarity 65.6%; Pred. No. 4.4e-111;
 Matches 292; Conservative 60; Mismatches 85; Indels 8; Gaps 5;

QY 20 LAVWGTLTPDV--GAVGKLEKEDLKRGFKLTDMAPLVAAEKGFPEDEGLFYQLEAQA 77
 DB 32 LLSW-LARFAPVVLAEVGEAEKEDLKRGFKLTDMAPLVAEKGFEDEGLFYVLEAQA 90
 QY 78 NKKVMDRVVNGELDSHMLAPAPLAASVGFGRKADIEVPFSMGFNNAITVSNIEIHWOM 137
 DB 91 NKKVLLDRVIDGELDGAHMLAGOPLGATIGITQAVITAFSMDLNGMALTISNDVWQOM 150
 QY 138 KRNIPLEG-GKRVHPKADYLKPVVEKYAEKGFPMAMTFPPAGSHNIKRLRYLAAGIN 196
 DB 151 KENLAQSDGKPVHPKADALKPVVTSYRDQKAFNMGWFPVSTHNYELRWLAAGIH 210
 QY 197 PGYSPPO-DISGOIGADALLSTPPOMPSTLEAGTIFGTCVGEPMNQOAVRKGIVPV 255
 DB 211 PGYAAHKDNGSQINADVLLSTPPOMPATWEAGTIKGYCVGEPMNQOAVRKGIVPV 270
 QY 256 ITDEBLMKDPEKVFQVTKQMAEKYPTNYLAATKALIRAAIWLADNNKRRKALBMLAQ 315
 DB 271 VTDYELTKNNPEKVFQVADMAEKYPTNHIRVYKALIRAAHMLDEDDNNRTRAVVLSR 330
 QY 316 KQYVADVEVLAASNGTFFYEKDDKRALPDENTFFRHGASYPSSAAVYLTQLRRWGM 375
 DB 331 SEYVGADAEVIANSMGTFFYEKDKRQVDFNVFRYNAATPYYSDAIMWLTQMRWGM 390
 QY 376 INEFKDDNMYLDTAKVYRPDIYLAAKELVAGKAKADFP--ADTSIKPSQNFIDK 432
 DB 391 IEHQKDSWYMDIAKQVYRPDIYORAAEALIEGTLASDFPDFAKESGFRPQTHFIDQ 450
 QY 433 VPEFDANKPNDYAKAIGIKKQTV 457
 DB 451 IHYDGRSPNAYLQGFALGLKGSSEL 475

RESULT 3
 Q8D6X8 PRELIMINARY; PRT; 487 AA.
 AC Q8D6X8;

DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DE ABC-type nitrate/sulfonate/bicarbonate transport system, periplasmic component.
 GN OrderedLocusNames=VW20393;
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxId=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAN-CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016809; AA007349.1; -.
 KW Complete proteome.
 SQ SEQUENCE 487 AA; 54515 MW; PESD5F11C5B865F4 CRC64;

Query Match 64.3%; Score 1572; DB 2; Length 487;
 Best Local Similarity 65.1%; Pred. No. 5.2e-110;
 Matches 289; Conservative 63; Mismatches 86; Indels 6; Gaps 4;

QY 20 LAVWGTLTPDVA-VGKLEKEDLKRGFKLTDMAPLVAEKGFEDEGLFYQLEAQA 78
 DB 44 LLSWLAFAAPVVLAEVGEAEKEDLKRGFKLTDMAPLVAEKGFEDEGLFYVLEAQA 103
 QY 79 KVVVMDRVVNGELDSHMLAPAPLAASVGFGRKADIEVPFSMGFNNAITVSNIEIHWOM 138
 DB 104 KVVLLDRVIDGELDGAHMLAGOPLGATIGITQAVITAFSMDLNGMALTISNDVWQOM 163
 QY 139 PNIPLEG-GKRVHPKADYLKPVVEKYAEKGFPMAMTFPPAGSHNIKRLRYLAAGIN 197
 DB 164 PNLVSDGKPVHPKADALKPVVTSYRDQKAFNMGWFPVSTHNYELRWLAAGIH 223
 QY 198 GYSPPO-DISGOIGADALLSTPPOMPSTLEAGTIFGTCVGEPMNQOAVRKGIVPV 256
 DB 224 GYVAPHKDNGSQINADVLLSTPPOMPATWEAGTIKGYCVGEPMNQOAVRKGIVPV 283
 QY 257 TDEBLMKDPEKVFQVTKQMAEKYPTNYLAATKALIRAAIWLADNNKRRKALBMLAQ 316
 DB 284 TDEYELTKNNPEKVFQVADMAEKYPTNHIRVYKALIRAAHMLDEDDNNRRAEYVLSRS 343
 QY 317 QYVADVEVLAASNGTFFYEKDDKRALPDENTFFRHGASYPSSAAVYLTQLRRWGM 376
 DB 344 EYVGADAEVIANSMGTFFYEKDKRQVDFNVFRYNAATPYYSDAIMWLTQMRWGM 403
 QY 377 NEFKDDNMYLDTAKVYRPDIYLAAKELVAGKAKADFP--ADTSIKPSQNFIDK 433
 DB 404 EHOKPDSWYMDIAKQVYRPDIYORAAEALIEGTLASDFPDFAKESGFRPQTHFIDQ 463
 QY 434 PEDANKPNDYAKAIGIKKQTV 457
 DB 464 RYDGRSPNAYLQGFALGLKGSSEL 487

RESULT 4

Q87HA7 PRELIMINARY; PRT; 454 AA.
 AC Q87HA7;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DE Putative nitrate transport protein.
 GN OrderedLocusNames=VPA1058;
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxId=670;
 RN [1]
 RP SEQUENCE FROM N.A.

CC -1- FUNCTION: Probably part of a high-affinity binding protein-dependent transport system for nitrate. Probably responsible for energy coupling to the transport system.

CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).

CC -1- FUNCTION: Probably part of a high-affinity binding protein-dependent transport system for nitrate. Probably responsible for energy coupling to the transport system.

CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).

```

CC -1- SIMILARITY: Belongs to the ABC transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; D90906; BAA17490.1; -.
DR PIR; S77387; S77387.
DR HSSP; OS8206; 1L2T.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_Transporter.
DR InterPro; IPR005890; NtrCD.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfam; TIGR01184; ntrCD; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR KMW; ATP-binding; Complete proteome; Membrane; Nitrate assimilation;
DR Transport.
DR MAIN; 1 254 ABC transporter.
FT DOMAIN 255 278 Linker.
FT DOMAIN 279 670 NtrA-like.
FT NP_BIND 42 49 ATP (potential).
SQ SEQUENCE 670 AA; 75100 MW; 03B47B6C7918AD14 CRC64;

Query Match 26.6%; Score 651; DB 1; Length 670;
Best Local Similarity 36.3%; Pred. No. 2.3e-40;
Matches 158; Conservative 68; Mismatches 153; Indels 56; Gaps 13;

QY 28 ADVGAVGK-----LEKEDLKFGFIKLTDMAPLAVAEKGFEDEGL-FVQLEAQA 77
DB 260 AKKGVAVSGPFAEMGNGLEKINDLGFIPLDCAPLVAKEGFGQKGLBEVNLVKEP 319
QY 78 NKKVMDRVVNGELGDSHMLAPLAASVGFGRADIEV--PSSWGFNGNAITVSNELWH 135
DB 320 SMOAIDGIRERRLDGAQWAGWPLALTIGMGSKTPLPWTVAMVMSRNGNAITLSKKFAB 379
QY 136 QMKPNPLBEGKGVHPIKADYLKPVVEKKAEKPPNMAMTPPASGSHNKLRYWLAAGI 195
DB 380 -----AGVKTLEDLR-----LKLATPDQVSTLGVHPSNQNLLRLRWLASGSI 424
QY 196 NPGVYSPPODISQIGADALLSVTPPOMSTLEAGTIFGYCVGEPMNQOAVFKGIGVPE 255
DB 425 D-----PDQDIN-----LWRLPPQVNSNLEAGNIDGFCVGEPMNSYAAVKNLGYVI 471
QY 256 ITDEELMKDTPPEKVFQVTKQMAEKYPNTYLAATKALIRAAIWLADNNKRNKEAIEMLAQ 315
DB 472 ATCDLDMNGHPEKVLDMREBWKVPATHLALVKALLEACEY--CDRRHRRQIILDYAL 529
QY 316 KQYGVADVLAASNMGTFEYEKD-DKRALPDENTFRRGASAPSSASVAVWLTOLRRNG 374
DB 530 PQYVGSTIYISPGFITEYDQGNDAEMLLDENQYVQSNPSRSEGLMILTOLARWG 589
QY 375 MINEFKPDWMLDTAKNVRPDIYLAAXELVAEGAKAEDPADPISIKPSQNFIDKVP 434
DB 590 YID--PPKKW-VBIIRVRRPDLFGAACHL--GMPDLEGDHNVSL-----PDGAV 636
QY 435 PDANKENDYLAAPAI 449
DB 637 FFPNDPLGYIKRFTI 651

RESULT 7
Q8U925 PRELIMINARY; PRT; 430 AA.
AC Q8U925;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

```

```

DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ABC transporter, substrate binding protein.
GN Name=ntcA; OrderedLocustNames=Atu3905;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Serubal J.C., Kaul R., Monks D.E., Ktajiama J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Cleedenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
DR EMBL; AB009321; AAL44713.1; -.
DR PIR; AC0307; AC0307.
DR PIR; H98248; H98248.
DR TIGRfam; TIGR01409; TAT_signal_seq; 1.
DR Complete proteome.
SQ SEQUENCE 430 AA; 46786 MW; 8388162BDE02B2D8C CRC64;

Query Match 25.7%; Score 627.5; DB 2; Length 430;
Best Local Similarity 35.5%; Pred. No. 7.7e-39;
Matches 166; Conservative 70; Mismatches 171; Indels 61; Gaps 19;

QY 1 MNTIIRSS-SKKULL--TISASLAVWGLTIAPD--VGAQKLEKEDLKFGFIKLTDMAPL 55
DB 1 MKKIFSGIVSRRTILKTTATYALVTAVRTAPSGAPAAAEPAEPAKAGIPLALTDAAPL 60
QY 56 AVAAEKGFEDEGL-FVQLEAQAANKVMDRVVNG-----ELDGSHMLAPLAASVGFGR 110
DB 61 IIAEKGFLPAKKGMPVEVLKQASWCATRDNLVLGASNSIGIDGAILTTPYLMHTGKYT 120
QY 111 KADIEVPS--MGFNNAITVSNELHQMKNPIPLBEGKGVHPIKADYLKPVVEKKYA 166
DB 121 QNNVPEVPTILARLINDSGISVAKK--YAETGVOLDASK-----LKAAEFKKKA 168
QY 167 EKGPPNMAMTPPAGSNIKLRYWLAAGINPGYSPPODISQIGADALLSVTPPOMPS 226
DB 169 DGEIKRAMTTPGGTHDLWIRYLLAAGID-----PDQVNS-----TIVPPQOMVA 215
QY 227 TLEAGTIFGYCVGEPMNQOAVFKGIGVPIITDEELMKDTPPEKVFQVTKQMAEKYPNTYLA 286
DB 216 NKKVGMDFCVGEPMNEBOLVNOGIGFTACTTGELMKHPEKALGRADWVENKPNPATGA 275
QY 287 VTKALIRAAIWLADNNKRNKEAIEMLAQKQYGVADVLAASNMGTFEYEKDDKRALPD 346
DB 276 LLMAVEAQAQW--CEBMANKEEMSTILGRQMFNPVPPKQVLGRLLKGNINY--GNGRVLE- 330
QY 347 FNT-----PFRGASAPSSASVAVWLTOLRRGMINEFKEDWMLYDTAKNVRPDIYLA 401
DB 331 -NTGLQMKFWQDHASVPPFHSWFTTENIRNG--KRAPDIDYALVAKVVRREDIWRRA 386
QY 402 AKELVAEGAKAEDPADPISIKPSQNFIDKVPFDANKENDYLAAPAI 449
DB 387 AKDL---GVA---DLPASTS-RGKETFPDGKY-FDPENPSATLBSLSI 426

RESULT 8
Q7CTL1 PRELIMINARY; PRT; 471 AA.
AC Q7CTL1;
DT 05-JUL-2004 (TRENBLrel. 27, Created)

```

DT 05-UTL-2004 (TReMBLrel. 27, last sequence update)
 DT 05-UTL-2004 (TReMBLrel. 27, last annotation update)
 DE AGR_L_1886p.
 GN OrderedLocuNames=AGR_L_1886;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NCBI_TaxID=176299;
 RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
 RC STRAIN=Cereon.
 RA Goodier B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Guetillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
 RA Hounell K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
 RA Mollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AEO08293; AAK89514.1; -
 DR InterPro; IPR009073; HSC20_C.
 DR TrIRFAMS; TIR01409; TAT_signal_seq; 1.
 DR TrIRFAMS; TIR01409; TAT_signal_seq; 1.
 SQ SEQUENCE 471 AA; 51726 MW; 2650A5968C39D79A CRC64;
 Query Match 25.7%; Score 627.5; DB 2; Length 471;
 Best Local Similarity 35.5%; Pred. No. 8.7e-39;
 Matches 166; Conservative 70; Mismatches 171; Indels 61; Gaps 19;
 QY 1 MKTIIRSS-SKLLT--TSLASLAWGLTIAPD--VGAVGLEKEDLKFGFIKLTMDAPL 55
 DB 42 MKKIFGTVSRRIKLTATLVAALVTRAPFSGAFAPATAPREVKAKGFIALTLDAFL 101
 QY 56 AVAAEKGPEDEGL-FVQLEAQNKKYVMDRVNG---ELDSHMLAPPLAASVGFCT 110
 DB 102 IIAAEKGLPAKGMPEVEVLKQASMGATRDNLVGGASNGIDGAILITPMPYLMHTGKVT 161
 QY 111 KADIEVPFS---MGFNMAITVSNELWQMKNPLEGSKPHPIKADVLKRVKYYKA 166
 DB 162 QNNVPFMTILARLNDISQISVAK--YAETGVQIDASK-----LKAPEKKKA 209
 QY 167 EGKPFNNAMTPFAGSHNIKRLRYLWLAAGINPGYSPPODISQIGADALLSYTPPQMP 226
 DB 210 DGEIIGAANTFPGCTHDWIRWYLAAGID---FDKQVS-----TIVPPQWVA 256
 QY 227 TIEAGTIFGTCGEPNNQAVFKGIVPVITDELMKOTPEKVFQVTKQMAEKYPNTYLA 286
 DB 257 NKKVGNMDFCVGEPPNNEQLVNGIGFTACTTGELMKGHEKALGMRADVVEKPNATKA 316
 QY 287 VTKALIRALWLDADNNKKEALIMLAQKQYGVADVEVLAASMGTFEYKDKALPD 346
 DB 317 LIMAIVEAQW--CDMANKEEMSTILGKQFNVPPKDLVGLKKNIN--GNGRVLE- 371
 QY 347 FNT-----FFRHGASYPSYSAVWYLTQLRRMGIMNEFKPDNMYLDTAKKVVPPDIYLA 401
 DB 372 -NTGLQMKWQDASYPFHSHDSMTENTIRWG---KFAPTDVYKALVAVNRREDIWRRA 427
 QY 402 AELVAEGKAKADEPADTSIKPSQNFIDKVPDANKPNDYLAAPAI 449
 DB 428 AKDL---GVA---DLPASTS-RGKETFPDQKV-FDBENPFAVLESISI 467
 RESULT 9
 Q92VK8 PRELIMINARY; PRT; 441 AA.
 AC Q92VK8;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE Putative nitrate transport protein.
 GN ORFNames=SMD21114;

OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymB.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 NCBI_TaxID=382;
 RX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
 RC STRAIN=1021;
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorholzer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouy J.,
 RA Golding B., Puehler A.,
 RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL591985; CAC49093.1; -
 DR TrIRFAMS; E95928; B95928.
 KW Complete proteome; Plasmid.
 SQ SEQUENCE 441 AA; 47208 MW; 2F2F4498BBF81D58 CRC64;
 Query Match 25.2%; Score 616.5; DB 2; Length 441;
 Best Local Similarity 33.8%; Pred. No. 5.4e-38;
 Matches 153; Conservative 74; Mismatches 182; Indels 43; Gaps 12;
 QY 19 SLAWGLTIAPDVGAVGLEKEDLKFGFIKLTMDAPLVAAEKGPEDEGLFVQLEAQN 78
 DB 19 NLSAGALPAP--GSIRHSGSRTRVAGFIPLVDAALVIAAEKGFQREGITELVQDVS 76
 QY 79 WKYVMRVNNGELDSHMLAPPLAASVGFCTADLEV--PESMGFNMAITVSNELWQ 137
 DB 77 WANVRRLAFRQPDVAHMLSPMVAAMLGSGNSPSTIAPFSGRGGNALTLSTRYGLM 136
 QY 138 KPNIPLEGKVPPIKADYLKPYVEKYKAEK--PFNNAMTPFAGSHNIKRLRYLWLAAGIN 196
 DB 137 QQAGGLGGEEDALK-NAKLAALVIRSAAGAARPLTLGTYTPSSINVERRYLAAGID 195
 QY 197 PGYSEPPDISQIGADALLSYTPPQMPSTIEAGTIFGTCGEPNNQAVFKGIVPVIT 256
 DB 196 P-----DSDVTLVVPVPPMTSDALAAGIDGFCVGAQNNMNVASERGVGRIVA 242
 QY 257 TDELMKOTPEKVFQVTKQMAEKYPNTYLAALVTKALIRALWLDADNNKKEALIMLAQK 316
 DB 243 TKQDIMPASPEKTYGNRPMAEARNRDTVSRLLVALDRAARW--SDPANNGRLLAEVLAER 300
 QY 317 QYGVADVEVLAASMGTFEYKED-DKRALPDENVTFRRHGASYSYSAVWYLTQLRRMG 375
 DB 301 RHVAAPADIIIRRYLAGEFALIDEGNRRVENVYVFAAGPANYRPPQALMTYSQMRWQ 360
 QY 376 INEFPKDNMYLDTAKKVVPPDIYLAAKELVAEGKAKADEPADTSI-----KPSQNF 430
 DB 361 TSLSRQ---RVDAISAVRPDLVREAL-----GK---DALPADADAGRLEGATDGRFM 408
 QY 431 DKVPDANKPNDYLAAPAIGLKQTVVAGSKV 462
 DB 409 DGHVPDPGRIDEDYIKSF-----GAPSVNRSRV 435
 RESULT 10
 Q8U924 PRELIMINARY; PRT; 431 AA.
 AC Q8U924; Q7CTL2;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, last annotation update)
 DE ABC transporter, nucleotide binding/ATPase protein
 DE (AGR_L_1881g1p).
 GN Name=ntC; OrderedLocuNames=AGR_L_1881g1; Atu3906;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NCBI_TaxID=176299;
 RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
 RC STRAIN=Cereon.
 RA Goodier B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Guetillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
 RA Hounell K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
 RA Mollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AEO08293; AAK89514.1; -
 DR InterPro; IPR009073; HSC20_C.
 DR TrIRFAMS; TIR01409; TAT_signal_seq; 1.
 DR TrIRFAMS; TIR01409; TAT_signal_seq; 1.
 SQ SEQUENCE 471 AA; 51726 MW; 2650A5968C39D79A CRC64;

```

RC STRAIN=DuPont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Glendinning T., Deatherage G., Giller W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphitachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Ioo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Gurolo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Hounmel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AEO09321; AAL44714.1; -
DR EMBL; AEO08293; AAK89513.1; -
DR PIR; AD3037; AD3037.
DR PIR; G98248; G98248.
KM Complete proteome.
SQ SEQUENCE 431 AA; 46254 MW; 22DB764D66992C63 CRC64;

Query Match 25.0%; Score 611.5; DB 2; Length 431;
Best Local Similarity 35.5%; Pred. No. 1.2e-37;
Matches 151; Conservative 69; Mismatches 166; Indels 39; Gaps 12;

QY 34 VGLKEKEDIKFGFKITDMAPLAVALAEKGFEEDEGLFVLEAQAQWVMDRVNGELDG 93
DB 24 VQSDROKIRAGFIPLVDASVILAAAFGPADEGLILDVQVSWANVRDLAFQFDI 83
QY 94 SHMLAPAPLAASVGFCTKADIEV-PSMGFNGAIVSNIEIHQMKPNPL-EGKPVHP 151
DB 84 AHMLSPMPVASMGLGNSNPSTPTTSPSLRGNAITLSTRLFARMKALTGLSETAGALEN 143
QY 152 IKADYLPVVEKKAEGK-PFNMAATFPAGSHNIKRLWLAAGINPGYSPQDISQOI 210
DB 144 ARA--LKLVLDMRANGAEPPLTGMYTFFSSSHYERFVYLAAGIHPD----- 190
QY 211 GADALISVTPPQMPSTLEAGTIFGVCGVBPNNQAVFKIGVGVITDELMKDPKVF 270
DB 191 --DVKLIVVVPPTMTSDALAAAGIDGFCVGAQWNIYVAERGVGRIVAQKODLWPSAEKYI 248
QY 271 GYTKQMAEKYKPTTYLAVTALTRAAATWLD-ADNNKRRKRAIEMLAOKOVGADVLAAS 329
DB 249 GMRPEWASQOEVEGLTALDPAASWCDLADNDLSCA---LADPRVYIAQOSIIRV 305
QY 330 NNGTFEYE-KDDKRALPDENTFFRHGASYSYSSAVWYLTOLRRWMINEFPDMWYLD 388
DB 306 LAGEPSIDQGNRRVIEKFTFHGDHANTFROSQISIMITYSMTKWC---QAEISGTGVA 362
QY 389 AKNVVRPDIYLAARELVAEGKAKADFPADTSIK-----PSQNFIDKVPFANKENDYL 444
DB 363 ALSAVRPDIYRAA-----LGDGKA-----PGDADIRIEGQDEGRFVDFVPADIAGTV 413
QY 445 AKFAI 449
DB 414 NSFAY 418

```

RESULT 11

```

NRTA_ANASP
ID NRTA_ANASP STANDARD; PRT; 440 AA.
AC Q44292; O06469;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nitrate transport protein nrtA.
GN Name=nrtA; OrderedLocustNames=aln0608;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
ON NCBI_TaxID=103690;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=97144534; PubMed=8990301;
RA Frias J.E., Flores E., Herrero A.;
RT "Nitrate assimilation gene cluster from the heterocyst-forming
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL J. Bacteriol. 179:477-486(1997).
RN
RP REVISIONS.
RA Frias J.E.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
RN
RP SEQUENCE OF 1-92 FROM N.A.
RX MEDLINE=97136629; PubMed=8982006;
RA Cai Y., Wolk C.P.;
RT "Nitrogen deprivation of Anabaena sp. strain PCC 7120 elicits rapid
RT utilization of a gene cluster that is essential for uptake and
RT utilization of nitrate.";
RL J. Bacteriol. 179:258-266(1997).
RN
RP FUNCTION: Essential component of the nitrate-transporting system.
CC May be the substrate-binding protein (By similarity).
CC -SUBCELLULAR LOCATION: Inner membrane-associated (potential).
CC -SIMILARITY: Strong, to carotenoid-binding protein A (cbpa).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X99709; CA68041.2; -
DR EMBL; AP003583; BAB72565.1; -
DR EMBL; U61496; AAC46075.1; -
DR PIR; AG1882; AG1882.
DR InterPro; IPR006311; Tat.
DR TIGRFAMs; TIGR01728; Ssua fam. 1.
DR TIGRFAMs; TIGR01409; Tat_signal_seq. 1.
KM Complete proteome; Inner membrane; Nitrate assimilation; Transport.
FT CONFLICT 100 W -> C (in Ref. 1).
SQ SEQUENCE 440 AA; 48475 MW; 29937411FB45CE9C CRC64;

```

```

Query Match 25.0%; Score 610.5; DB 1; Length 440;
Best Local Similarity 33.4%; Pred. No. 1.5e-37;
Matches 158; Conservative 71; Mismatches 173; Indels 71; Gaps 17;

QY 8 SSKLLLT-----LSASLAWGLTI-----APDVGAVG--KLEKEDLK 44
DB 5 SRKRLFTTGAAMAAASILVHGCTSGSOSATTGEQAPSAAPAAVYSAANAPVETTKAKL 64

```

```

QY 45 GFPLTDMAPLVAAEKGFPEDEGLF-VQLEAQAANKVMDRV---VNGELDGSHTLAP 99
DB 65 GFPLTDMAPLVAAEKGFPEDEGLF-VQLEAQAANKVMDRV---VNGELDGSHTLAP 124
QY 100 APILAASVGTGKDIIVPFGMGNGNAITVSNELIWHQMKNPILLEGSKPHPIKADVLKP 159
DB 125 MFLVMTI--NDKVPMTILARLNTNGAIVAEKF--KELNVNLE-----SSSLKD 170
QY 160 VVEKYAEKGFPMNMTFPAAGSHNIKLRVYLAAGINPGYSSPPODISQIGADALISVT 219
DB 171 AAIKAKADKKALQAGITFPGGTHDLMRWYLAAGINP-----DDDVYLEAV 217
QY 220 PPPQMPSTLEAGTIFGVCVGEPMNQAVFKGIGVPIVITDELMKDPPEKVPYTKQMAEK 279
DB 218 PPPQVAMNMKVNTVDGFCVGEPMNQAVLNQIGSALVITGELWKDHEKAFPSRWQDIEQ 277
QY 280 YPPTVYLAVTFTALIRAAIWLADANNKREKAIEMLAQKQYGVADVEVLAASMGTFEEKQ 339
DB 278 NPMQAQIILMAIIEAQOM--CDKAKENKEEMCKICSDKRYFNVAADIIERAKGNIDY--G 333
QY 340 DKRALPDF--NTFFRHGASYPSSAVMYLTOLRRKMGINEFKPDNMYLDTAKNYYRPD 396
DB 334 DGRKEQNFARMKFMADNMSYFYKSHDIWFLTEDIRKGYLPK---DTKQDYNQVNNKED 390
QY 397 IYLAALKELVAEKKAADEFPADTSIKPSQNFPIDKVPDPANKRNDYLAKPAI 449
DB 391 LMKKAAK--AIGVADAE-IPASS--RGVETFF-DGVKFPDEKPEEYINSLKI 437

RESULT 12
Q7NEC4 PRELIMINARY; PRT; 429 AA.
AC Q7NEC4;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DE G113956 protein.
GN OrderedlocusNames=g113956;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
NCBI_TaxID=33072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids."
RL DNA Res. 10:137-145 (2003).
DR EMBL; AP006581; BAC91897.1; -
KW Complete proteome.
SQ SEQUENCE 429 AA; 47153 MW; 205101A5C8B22F9 CRC64;

Query Match 25.0%; Score 610; DB 2; Length 429;
Best Local Similarity 35.4%; Pred. No. 1.6e-37;
Matches 162; Conservative 60; Mismatches 174; Indels 62; Gaps 14;

16 LSASLAVWGLTI-APDVGAVKLEKEDLKFGFIKLTDMAPLVAAEKGFPEDEGLFVQLE 74
DB 12 IAAATAITGFTLWPKYALAKGALKRPLKYGFIALIDCAPLVIAAREGFFERBGLDVELS 71
QY 75 AQANMKVMDRVVNGELDGSHTLAPILAASVGTGK--ADIEVPFGMGNGNAITVSN 132
DB 72 KESWSMASVREGGLITGRLDASHALAGTPLAVALGABGAPAPLITAMSLDINGNAITPSKR 131
QY 133 IWHQMKNPILLEGSKPHPIKADVLKPVKRYAEKGFPMNMTFPAAGSHNIKLRVYLA 192
DB 132 LM-----QAGVRSG-----AD-LKKIITGKY-GTTLTGAMYSASSMTNINLCYLAH 177
QY 193 GGINPGYSSPPODISQIGADALISVTPPPQMPSTLEAGTIFGVCVGEPMNQAVFKGIG 252

```

```

DB 178 HDIHP-----YRDVRLITLTPAQILANIEAGNIDFECVTEPNNTHAHEGAG 224
QY 253 VPVITDEELMKDTPPEKVPFGVTKQMAEKVPNTYLAATVTKALIRAAIWLADANNKREKAIEM 312
DB 225 FTVIADRDVWGHPKEVYLAAMEPMAITHPTHTLALKALLAEGRY--CDEPARHQSVLRT 282
QY 313 LAQKQYGVADVEVLAASMGTFEEKQD-----KRALPDENFTFRHASY----- 357
DB 283 LSLRTVLAQPLFLIEPALSGRLDFGFGDEGKATASLKDVPDEVVEFKRADYLVGNDAQ 342
QY 358 --PSYSSAWVYLTOLRRKMGINEFKPD--NMYLTPAKNVRPDIYLAALKELVAEKKAAE 414
DB 343 TFFWKSHGIMLTQMARWQIPAIIPAGVDKLD--RYRVVDIYRQAAELGI--KAPAQ 397
QY 415 DFPADTSIKPSQNFPIDKVPDPANKRNDYLAKPAIGLK 452
DB 398 DY-----KSENTFTIDRRRFPDSNVAYLDSFEIAR 428

RESULT 13
Q7NKB1 PRELIMINARY; PRT; 427 AA.
AC Q7NKB1;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DE Nitrate/nitrite transport system substrate-binding protein.
GN Name=ntrcA; OrderedlocusNames=g11567;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
NCBI_TaxID=33072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids."
RL DNA Res. 10:137-145 (2003).
DR EMBL; AP006573; BAC89508.1; -
DR InterPro; IPR006311; Tat.
DR TIGRFBMS; TIGR01409; TAT_signal_seg. 1.
KW Complete proteome.
SQ SEQUENCE 427 AA; 46542 MW; C0D86F1625E39273 CRC64;

Query Match 24.3%; Score 594; DB 2; Length 427;
Best Local Similarity 33.8%; Pred. No. 2.6e-36;
Matches 152; Conservative 73; Mismatches 169; Indels 56; Gaps 14;

9 SKLLITLSASLAVWGLTIAPDVGAVKLEKEDLKFGFIKLTDMAPLVAAEKGFPEDEG 68
DB 11 TRRGGLQAAAATAATVGSILTFPGYVGAADAPETTKARIGFISLSDCAPLVIAAREGGLFDKYG 70
QY 69 L-FVQLEAQAANKVMDRV---NSELDSHTLAPILAASVGTGKADIEVPF---S 119
DB 71 MKDVEAKQKQSWGVTDNLELGAAGGIGDGAHLLTPVYVLIANGNTKSGSKVPMFTLAR 130
QY 120 MGFNGNAITVSNELIWHQMKNPILLEGSKPHPIKADVLKPVKRYAEKGFPMNMTFPA 179
DB 131 LNVNGGIGISVANK--YKL-----LKYGLDAAPKAEALKR-----KANGDPITVAQTFPG 178
QY 180 GSHNIKLRVYLAAGINPGYSSPPODISQIGADALISVTPPPQMPSTLEAGTIFGCVG 239
DB 179 GTHWAMLRVYLAAGIDP-----ETDVKMITVPPQVAMNMKVNTVDGFCVGEPMNQAVFKGIG 225
QY 240 EPMNQAVFKGIGVPIVITDEELMKDTPPEKVPFGVTKQMAEKVPNTYLAATVTKALIRAAIWL 299
DB 226 EPMHQIILNQIIGYVAVTGTQIWNRRHEKSFALRADYVEKTPATATALLMAVOEADIW-- 283

```

```

OY 300 ADNNKRNKAIAEMLQKQYGVADVEVLAASMGTFEY--EKDQKRALPDGNTFFRRGASV 357
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 264 ADKANKKDELAQVSGSRGSIWGAIPVSDIVARIYKGIIDYGGRPVEPKRSPHIMQWKPAST 343
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
OY 358 PSYSAVWYLTQLRRKGMINEKRPDWNVLDTAx--NVYRPDIYLAARELVAEGRAKAE 414
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 344 PYQSHDLWLTEDIRHGVLPJ-----TDTTKLVAAVNRBDLMBRAAKL-----GQ 390
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
OY 415 DEPADTSIKPSQNFIDKVPFPDANKNDVL 444
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 391 PAPKQTS-RGVEKFP-DGVAFPPTKPEAVL 418
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 14
O8Y276 PRELIMINARY; PRT; 657 AA.
ID O8Y276
AC O8Y276;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nitrate transport ATP-binding protein.
GN Name=ntc; Ordered locus names=at10610;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxId=1033690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; Pubmed=11759840; Kuritz T., Sasamoto S.,
RA Kaneo T., Nakamura Y., Wolk C.P., Kawashima K., Kimura T.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Matsuno A., Muraki A.,
RA Kishida Y., Kohara M., Matsuno M., Takazawa M., Yamada M.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AP003583; BAB72568.1; -.
DR PIR; A11882; A11882.
DR HSP; Q58206; IL2T.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042628; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0015112; F:nitrate transporter activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0015706; F:nitrate transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR Pfam; PF00005; ABC_tran; 1.
DR Prodom; PD000006; ABC_transporter; 1.
DR SMART; SMO0382; AAA; 1.
DR TIGRFAMs; TIGR01184; ntrCD; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PSS0893; ABC_TRANSPORTER_2; 1.
DR ATP-binding; Complete proteome.
SQ SEQUENCE 657 AA; 73294 MW; F7E434798109324 CR664;

Query Match 24.0%; Score 587; DB 2; Length 657;
Best Local Similarity 35.8%; Pred. No. 1.5e-35;
Matches 150; Conservative 74; Mismatches 139; Indels 56; Gaps 17

OY 39 KEDLKFSGIKLTDMAPLVAAEKGFPEDEGL-FVQLEAQAANKVYMDRVYNELDQSHL 97
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 276 KAVIEIGMPLTDSAPLIVAEEKGFPAKGYGLDNLVILNRANNQALATGVVTKLDAAOVY 335
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
OY 98 AAPLAASVGFCTKADIEV--PFSMGFNGNALTIVNEIWHQKPNIPLEGKRPVHPKAD 155
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 336 AGMPIALTLGAGSQCPFTVYNALNLSRANALITPKRKLTYNGVRSL-----AD 383
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
OY 156 YLKPVEYKKAEGKPFNNAMTFPAGSHNIKLRYMLAAGINFGYSPPODISGQIGADAL 215
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 384 -LKKAVID--SSPDQILTLGVVSHASQMQLILRYMLAAGID-----PRDVS----- 427
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
OY 216 LSVTPPPQMPSTLGAQITFGYVGVGPANQAVFKGIQVPIVITDEELMKDTPKRGVTVK 275
      ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

```

D6      428  LTIIPFQWWSQLKAGNIDIGYCAQGEPMNYQAVHDDLGVAATTALEINSGQPKVLGTRED 487
QY      276  MAEKYENTYLAVALKALIRAAIWLDDNNKNRKEALEMAQKQYGVADVEYLAASMGTF- 334A
D6      488  MAQKYETETINLVKALIEACKY--CDDLNRNREIIEILCREY----LDVNPAYVRSGFT 544A
QY      335  -EYEKD---KRALPDFTNTEPRHGAASYPSYSSAAVYLTOLRRMGINEFKPDNNVLTDTAK 390A
D6      542  DPLDRDDGTPPPOOLTRYNGFYLNKTNINRNREILIMITQMARWG-LTFP-PKMW-VEITE 598A
QY      391  NNYRPDIYLAARKELVAEGKAEDFPADTSIKESQNEFIDKVPDANKPNDYLAKEAI 449A
D6      599  RVRKRTIIFGAARDL-----GLLDIGDDPI-----HLFDGKL-FNNSSEPIEYIKSEI 646A

```

RESULT 15	PRELIMINARY;	PRT;	420 AA.
ID 08Y2F5			
AC 08Y2F5			
DT 01-MAR-2002 (TREMBLERel. 20, Created)			
DT 01-MAR-2002 (TREMBLERel. 20, Last sequence update)			
DT 01-MAR-2002 (TREMBLERel. 20, Last annotation update)			
DE PUTATIVE NITRATE TRANSPORTER PROTEIN.			
GN Name=nasF, Synonyms=RS03349; OrderedAccessionNames=RS03081;			
OS Ralstonia solanacearum (Pseudomonas solanacearum).			
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;			
OC Burkholderiaceae; Ralstonia.			
OX NCBI_TaxID=305;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=GM11000;			
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;			
RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mengent S.,			
RA Ariat M., Billaud A., Brothier P., Camus J.C., Cattolico L.,			
RA Chardin M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,			
RA Siglier C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,			
RA Weissensbach J., Boucher C.A.,			
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.,"			
RL Nature 415:497-502(2002)			
DR EMBL; AL646058; CADD3909.1; -			
KM Complete proteome.			
SQ SEQUENCE 420 AA; 45516 MW; 6BC0DBAF8B2F3C CRC64;			
Query Match 24.0%; Score 585.5; DB 2; Length 420;			
Best Local Similarity 32.8%; Pred. No. 1.1e-35;			
Matches 151; Conservative 82; Mismatches 164; Indels 63; Gaps 15			
QY 3 TIRSSSKLLITLS--ASLAWGLITPAVDYGVNGK--LEEKEDIKFGFTIKTDMAPLAVA 58			
DB 9 TPIPKRRRYLATVATAGGSAMALIDPLVRAAGAMAGSDAPEKTEIKVGIPLTDCASVMA 68			
QY 59 AEKGFPEDEGLFVLEAQAWMKVMDRVNNGELGSHMLAPAPLAASGVF--GTRADIEVP 117			
DB 69 ATLGDKKYGKIKIYPSKEASWAGRDGLVSGDDLAAHVLGLYGVGLIGIGPKKDMAYL 128			
QY 118 FSWGFNGNAITVSNELIWHQMKPNIPLEGKRPVPIKADYIKPVVEKKYKAEKPFNMATF 177			
DB 129 MTLNNNOQAIIITLSKLT-----KEAGVRGDASLTALMTR--EKRDYTFADQF 172			
QY 178 PAGSHNIKLTWYMLAAGINGPYGYSPPDDISQIADADLLSTPPQPMOSTEAGTIPGYC 237			
DB 173 PYGTHAMWLYYMLAHGHIHP-----IQDAKAITVPPQVYANRKGVMDSYC 219			
QY 238 VGEPMNQAAVEFKGIGVITIDDELIMKDTPEKVFQVTKQMAEKYENTYLAATKALIRAIW 297			
DB 220 VGEPMQARRAIDAGIGFAETQAIWKDHPEKVLGTETAEFAQKYENTARALTAALVLEASKF 279			
QY 298 LDDANNKRRKEALEMLAKOKYGVADVVAASNMGTPEY-----EKDDGRALPDENFT 350			
DB 280 IDA--SASNRKRTAETVAASKYVNTDMDIILDRMIGRTNGIKGTWDDADPMR-----F 331			

3 409
- 381

This Page Blank (uspto)